Investigations into the mechanisms behind human imprinting and imprinting related disorders in model systems

Gareth Pollin
BSc (Hons) Biomedical Sciences
with Diploma in Industrial Studies

Research conducted within
School of Biomedical Sciences
Faculty of Life and Health Sciences
Ulster University

A thesis submitted for the degree of
Doctor of Philosophy

December 2020

I confirm that the word count of this thesis is less than 100,000 words.

Contents

Acknowledgements

Firstly, I would like to thank my supervisors, Professor Colum Walsh, for your continued support throughout my PhD. The experience and training I gained from you and Dr Rachelle Irwin will be the core foundations to my future career and for that, I could not be more grateful. I am also extremely thankful for the support and guidance throughout my PhD, I could not have asked for a better supervisor.

I would also like to acknowledge the rest of the members from the Walsh lab that have been with me throughout my time in the lab. Completing a PhD is a hard task for anyone, however, completing it while being surrounded by friends made it a much more enjoyable experience.

To my Grandfather Harry Gorman, you have always been a big influence and a great role model to inspire me to do my best. My Brother Kenneth, you have always been there when I needed you, from daily advice to sharing your Netflix, so thanks for everything. Dad thanks for always listening to me go on about my thesis, I look forward to having coffee meetups in the future. To the rest of my family, I would like to say thanks for their continuous support.

Roisin, thanks for being so understanding and patient while I was completing my thesis. I just want to let you know I appreciate everything you have done for me to ensure that I could focus on finishing my thesis and being so supportive of my career.

Saving the best for last. Mum, words could not describe everything you have done for me. Simply, I would not be the person I am today if it were not for you and for that, I will be forever thankful.

Thank you to all…

Abstract

Genomic imprinting plays a vital role in normal mammalian growth and development. Normal expression of imprinted genes is highly dependent on parental-specific methylation at the differentially-methylated regions known as imprinting control regions. Imprinting control regions often regulate the parental-specific expression of a number of nearby imprinted genes. Abnormal expression of these imprints is caused by chromosomal anomaly or an epimutation such as loss of methylation at the imprinted control region. Using human and mouse model systems I investigated the mechanisms behind imprinting and imprinting- like disorders.

Firstly, I utilised the chromosomally stable human fibroblast cell line hTERT-1604 with a long-term depletion of the maintenance methylation enzymes DNMT1 and UHRF1. I present data showing that the long-term knockdown of DNMT1 results in an interesting cross-talk between the gametic and somatic DMRs at a subset of imprinted clusters. Further to this, I show the depletion of UHRF1 results in loss of imprinting with the inability to recover upon the rescue of the WT protein.

I also took the opportunity to use publicly available datasets and a bespoke in-house Galaxy workflow to score abnormal methylation variability across patient samples and described a methylation index to detect epimutations in imprinting disorders. Not only was it able to detect the same epimutations as the wet-lab technique that was previously used to diagnose the patient samples, but the methylation index also detected epimutations in patient samples that were not previously diagnosed at the molecular level.

Finally, with the help of CRISPR-generated mouse models, I was able to describe a novel role for the PHD domain of Uhrf1 in the maintenance of genomic imprinting during embryonic development. Furthermore, I also contributed to the literature in support of Uhrf1's role in the repression of transposable elements.

Abbreviations

AA	Accelerated Ageing
ADCA-DN	Autosomal Dominant Cerebellar Ataxia, Deafness And Narcolepsy (ADCA-
aESC	Androgenetic ESC
AML	Acute Myeloid Leukemia
APOE	Apolipoprotein E
ART	Assisted Reproduction Technologies
AZA	5-Aza-2'-Deoxycytidine
BWS	Beckwith-Wiedemann Syndrome
CNV	Copy Number Variance
LOM	Loss Of Methylation
GOM	Gain Of Methylation
CRC	Colorectal Cancer
DKO	Double KO
DMRs	Differential Methylated Region
DNA	Deoxyribonucleic Acid
DNAmAge	DNA Methylation Age
DNMT	DNA Methyltransferase
DNMTi	DNMT1 Inhibitors
E	Embryo Day
ENA	European Nucleotide Archive
EPIC	Infinium® Methylationepic Beadchip
ERV	Endogenous Retrovirus
ESC	Embryonic Stem Cell
EWAS	Epigenome-wide Association Studies
FACS	Fluorescence-Activated Cell Sorting
FST	Fa/Scc/Tween
gDMR	Gametic DMR
GEO	Gene Expression Omnibus
GWAS	Genome-Wide Association Studies
HDAC1	Histone Deacetylase 1
HERVs	Human Endogenous Retroviruses
HET	Heterozygotic
HOM	Homozygotic
HSAN1E	Hereditary Sensory Neuropathy with Dementia And Hearing Loss Type 1E
IC	Imprinting Centre
ID	Imprinting Disorders
IFNs	Interferons
IQR	Interquartile Range
ISGs	Interferon-Stimulated Genes
KD	Knock Down
KO	Knock Out
LB	Luria-Bertani

LINE 1	Long Interspersed Nuclear Element-1
lncRNA	Long Non-Coding RNA
LTR	Long Terminal Repeat
MD	Microcephalic Dwarfism
MeCP2	Methyl-CpG Binding Protein 2
mESC	Mouse Embryonic Stem Cells
MI	Methylation Index
MLID	Multilocus Imprinting Disturbances
MS-MLPA	Methylation-Specific Multiplex Ligation-Dependent Probe Amplification
mUPD	Maternal UPD
MWU	Mann Whitney U Test
OMIM	Online Mendelian Inheritance In Man
ORF	Open Reading Frames
PAR	Pseudoautosomal Regions
PBR	Polybasic Region
PBS	Phosphate Buffer Saline
PBT	PBS/Tween
PCNA	Proliferating Cell Nuclear Antigen
PGCs	Primordial Germ Cells
PHD	Plant Homeodomain
phESC	Parthenogenetic ESCs
PHP1b	Pseudohypoparathyroidism Ib
PRC2	Polycomb Repressive Complex 2
PrE	Primitive Endoderm
pUPD	Paternal UPD
RFTS	Replication Foci Targeting Sequence
RING	Really Interesting New Gene
RNA	Ribonucleic Acid
T2DM	Type 2 Diabetes
TBRS	Tatton-Brown-Rahman Syndrome
TBST	Tris-Buffered/Saline/Tween
TDG	Thymine DNA Glycosylase Pathway
TE	Trophectoderm
TET	Ten-Eleven Translocation
TG	Thymine Guanine (TG)
TKO	Triple KO
TNDM	Transient Neonatal Diabetes Mellitus
TRD	Target-Recognition Domain
TSS	Transcription Start Site
TTD	Tandem Tudor Domain
UBL	Ubiquitin Like
UCSC	University Of California, Santa Cruz
uDMR	Unknown DMR

UHRF1	Ubiquitin Like With Phd And Ring Finger Domains 1
UPD	Uniparental Disomy
SAM	S-Adenosyl Methionine
UTR	Untranslated Region
WGBS	Whole-Genome Bisulphite Sequencing
WT	Wild Type
Xist	X Inactive-Specific Transcript
ZFP	Zinc Finger Proteins
ZGA	Zygote Genome Activation

Note on access to contents

I hereby declare that with effect from the date on which the thesis is deposited in the Research Office of the University of Ulster, I permit;

1. the Librarian of the University to allow the thesis to be copied in whole or in part without reference to me on the understanding that such authority applies to the provision of single copies made for study purposes or for inclusion within the stock of another library.

2. the thesis to be made available through the Ulster Institutional Repository and/or EThOS under the terms of the Ulster eTheses Deposit Agreement which I have signed.

Chapter 1- General Introduction

Genomic imprinting is the epigenetic phenomenon in which genes are expressed in a parent of origin manner. Epigenetics is the term given to the study of heritable alterations to genomic DNA that affect expression without changing the DNA sequence. These epigenetic regulators include chromatin and histone modifications and DNA methylation. While some sources include long non-coding RNAs (lncRNA) and microRNAs, these act at post-transcriptional levels and are not heritable through cell division and so are not regarded as epigenetic controllers by most researchers. The parental-specific expression seen at imprinted loci is largely controlled by the presence of imprinting control regions. These regions are differentially methylated regions that have been shown to have a functional effect. As the differentially methylated regions play an important role in the regulation of imprinted genes, this thesis will focus on the mechanistic action behind these regions. However, in this introduction, I will cover the basics behind methylation such as the establishment and maintenance of methylation as well as how aberrant methylation patterns can lead to imprinting disorders. This thesis will be investigating the changes of methylation seen at Differentially methylated regions (DMRs) with different models utilizing wet-lab techniques and publicly available datasets. Understanding the mechanisms that drive methylation at these DMRs and how they react to different models will help to understand the imprinted disorders related to these DMRs.

1.1 Addition, maintenance and loss of DNA methylation patterns

DNA methylation is the main contributor to the parental specific expression seen in imprinted genes through silencing or expression of imprinted genes through methylation or lack of methylation at differential methylated regions (DMRs) (Barlow and Bartolomei, 2014). DNA methylation was first discovered in 1925 in tubercle bacillus, though it was not until 1975 when it was found to regulate gene expression (Holliday and Pugh, 1975; Johnson and Coghill, 1925). Methylation is the addition of a methyl group to the 5’ carbon, primarily at a cytosine and guanine dinucleotide (CG). The methyl group is transferred to the 5’ carbon position of the cytosine base from the universal donor S-adenosyl methionine (SAM) (Gibney and Nolan, 2010). The human genome is quite CG sparse, however, where they are present they tend to

be methylated (Larsen et al., 1992). It is believed that the lack of CGs in the human genome could be credited to the inefficiency of the base excision repair (BER) pathway (Walsh and Xu, 2006). This comes about due to the hypermutability of CG sites in which the methylcytosine can deaminate to thymine resulting in a thymine guanine (TG) dinucleotide (Bird, 1980; Coulondre et al., 1978). When the TG mutation is not repaired this will transition into an adenine guanine dinucleotide and lead to loss of the CG site (Walsh and Xu, 2006). This is thought to have led, over time, to the loss of CG dinucleotides in most of the genome of organisms with cytosine methylation, but the preservation of CpG islands, where the cytosine is never methylated. Though CG methylation is the most predominant and most investigated form of methylation, it is important to note non-CG methylation occurs in some abundance in stem cells, oocytes and brain cells, though it can be found in low quantity in other somatic tissues (He and Ecker, 2015). Non-CpG methylation includes methylation at the cytosine of the following dinucleotides CpA, CpT and CpC, of which CpA is the most abundant (Jang et al., 2017). Interestingly, non-CpG methylation shows a spatial correlation with CpG methylation, while also showing a strong correlation with the expression of the *de novo* DNA methyltransferase (DNMT) DNMT3A (Ziller et al., 2011).

Methylation at the CG sites is the most abundant, though non-CG methylation also exists at a lower amount. However, non-CG methylation is not fully clear and is highly debated in the literature, some studies showed that non-CG methylation is relatively low at promoter regions and that increased levels of non-CG methylation results in repression of the gene. Furthering this increased levels of non-CG methylation at the promoter of peroxisome proliferator-activated receptor γ coactivator 1α gene was seen in type 2 diabetes patients which were also seen to have lower levels of expression of the gene (Barrès et al., 2009). However, it may not only be non-CG methylation at promoter regions that play a functional role as it was observed that methyl-CpG binding protein 2 (MeCP2) a protein involved in the neurological disorder Rett syndrome, was able to bind to gene body non-CG methylation. Investigating this, a study observed that mutations in MeCP2 resulted in misregulation of non-CG methylation at genes known to acquire increased levels of non-CG methylation after birth in a mouse model (Chen et al., 2015).

1.1.1 Methylation Distribution

The term CG island (CGI) is used for a region of at least 200 bases that is dense with CG sites (at least 50% CG rich) (Bird, 1986; Gardiner-Garden and Frommer, 1987). The regions within 2kb on either side of a CGI are termed as CpG "Shores", while the region that extends from 2kb upstream and downstream from the shore is termed the CpG "Shelf", while finally, any CpGs found anywhere else in the genome are classed as being in "open sea" (Borchiellini et al., 2019; Irizarry et al., 2009). Around 60% of promoter regions at genes within the human genome have a CGI present (Costello and Plass, 2001). CG islands are usually hypomethylated apart from about 3% of CGIs that are hypermethylated during development resulting in repression of the genes in which they are present (Illingworth et al., 2010). The hypermethylation at these genes is important for regulation during development and includes genes on the inactive X chromosome, imprinted genes and a subset of tissue-specific genes (Jones, 2012). Despite being obvious target regions for methylation, CGIs appear to be protected from methylation. At imprinted loci, histone modifications and transcription factors are believed to be a possible reason why these subsets of CGIs remain unmethylated. One well-investigated transcription factor that is associated with methylation is CTCF, which binds to unmethylated regions. Studies have shown that the loss of CTCF leads to hypermethylation at CGIs that are normally unmethylated (Engel et al., 2006; Schoenherr et al., 2003). Methylation of H3K4 is frequently observed at regions that facilitate transcription, for example, H3K4me3 has been shown to protect from methylation as it has been observed in at CGIs that are known to remain unmethylated in sperm (Bernstein et al., 2002; Henckel et al., 2012). Due to the nature of CGIs, they are effective regions to study for gene regulation in specific cell lineages and epigenetic biomarkers, while CpG shelves and shores are used for environmental effects on DNA methylation due to their susceptibility to external factors (Borchiellini et al., 2019).

Methylation found at the promoter region of a gene is mostly consistent with the subsequent repression of the gene (Klose and Bird, 2006). However, gene body methylation which is usually found on the exons of long transcripts away from 5' and 3' ends is commonly found at constitutively expressed genes such as housekeeping genes and rarely found at genes with variable expression (Zilberman, 2017). Maunakea *et al.*, (2010) showed that gene body

methylation was important for regulating intragenic promoters of the SHANK3 locus and its mouse homologue, suggesting a role of gene body methylation in repressing intragenic promoter expression. There is a clear contradiction between gene body methylation and promoter methylation which is referred to as the methylation paradox (Jones, 1999). Jjingo *et al.*, (2012) contradicting the study by Maunakea *et al.*, (2010) showed the relationship between transcription and gene body methylation is non-monotonic and that role of intragenic promoter repression was an epiphenomenal by-product that cannot explain the majority of methylation found at gene bodies and that gene-body methylation is due to the accessibility of DNA during transcription. The contradiction between gene body and promoter methylation is still not fully understood, however, the current consensus suggests gene body methylation facilitates gene transcription, as reductions are associated with concomitant decreases in transcript level (Irwin et al., 2014; Maunakea et al., 2010). Research on gene body methylation is important due to the association between aberrant methylation at gene-body in a variety of cancers (Arechederra et al., 2018).

Distal control regions, better known as enhancers, are known regulatory regions that are established to control gene expression outside gene promoters and were first discovered in 1981 (Banerji et al., 1981). Enhancers are known to regulate genes through their interaction with promoters at various distances due to chromatin looping, allowing them to be in close spatial proximity. In line with gene body methylation, enhancers correspond with cell-specific expression. (Ordoñez et al., 2019). While discovered around 40 years ago, there is still a lot of movement in this area with recent investigations describing coalescence of genes and several regulatory elements and super-enhancers which are characterised by their high occupancy for transcription factors (Schoenfelder and Fraser, 2019). This cell-specific gene expression regulation seen with enhancers is partly due to epigenetic regulation, with hypomethylation at the enhancers resulting in transcription factor binding and subsequent expression while hypermethylation is a common observation at inactive enhancers (Wiench et al., 2011). DNA methylation is not the only epigenetic regulator known to regulate enhancers, for example, enhancers are enriched with H3K27ac and H3K4me1 modifications which are associated with chromatin accessibility (Sur and Taipale, 2016). Epigenetic regulation at enhancers is observed throughout embryonic development and cell differentiation (Kozlenkov et al., 2014;

Meissner et al., 2008; Rönnerblad et al., 2014). Aberrant methylation at enhancers contributes to cancer progression and patient mortality (Bell et al., 2016).

Another control region found in mammals is an imprinting control region (ICR). ICRs are differentially methylated regions (DMRs) usually found at imprinting clusters that are seen to control allele-specific gene expression of several genes in a large region. The establishment of the imprinting mark is well-reviewed by Hanna and Kelsey (2014). ICRs control a subset of genes known as imprinted genes, they carry out this function due to a parental-specific methylation pattern leading to monoallelic expression (Edwards and Ferguson-Smith, 2007). Interestingly the DMRs that are maternally methylated are usually found at promoter regions while paternally methylated DMRs are usually found at intragenic regions (Bartolomei and Ferguson-Smith, 2011). The dosage control of imprinted genes carried out by the ICR is essential for many biological functions such as endocrine and metabolic functions, placental development and function, brain development and behaviour and foetal growth (Farhadova et al., 2019). Though methylation is the main driving factor of parental specific transcription it should be noted that histone modifications also show consistency with DMR methylation. H3 lysine-64 tri-methylation (H3K64me3), histone H3 lysine-9 trimethylation (H3K9me3), H4 lysine-20 trimethylation (H4K20me3) and H4 arginine-3 symmetrical dimethylation (H4R3me2s) is common at the DNA methylated allele, whereas enrichment of H3 lysine-4 di- and/or trimethylation (H3K4me2/H3K4me3) and acetylation of H3 and H4 at lysine residues is consistent with the unmethylated allele (Sanli and Feil, 2015)- these will be discussed in more detail later in this section.

1.1.2 De novo DNA methylation enzymes

DNA methyltransferases (DNMT) DNMT3A and DNMT3B are important regulators of *de novo* methylation. These were first discovered in mouse embryonic stem cells (mESC), which continued to methylate invading viral sequences even with the loss of DNA Methyltransferase 1 (DNMT1) (Lei et al., 1996; Okano et al., 1998). DNMT3A and DNMT3B regulate *de novo* methylation during gametogenesis and embryonic development, in which unmethylated CG sites become methylated. *De novo* methylation during embryonic development is not well understood, it is known that *de novo* methylation is attracted to repeated sequences though

it has currently not been fully explained why (Edwards et al., 2017; Garrick et al., 1998). Growing oocytes and spermatogonia express DNMT3L which is a co-factor that forms a complex with DNMT3A and DNMT3L, the addition of DNMT3L allows DNMT3A and DNMT3B to bind to H3K4me0 and subsequently results in *de novo* methylation (Bourc'his and Bestor, 2004; Bourc'his et al., 2001; Ooi et al., 2007). DNMT3A was also shown to be important for establishing methylation at non-CpG sites such as cytosine, thymine (CpT) and cytosine, adenine (CpA) dinucleotides (Ramsahoye et al., 2000). Dnmt3c is a *de novo* methyltransferase that is only found in mouse which protects the male germline from transposon activity by methylating evolutionarily young retrotransposons and is required for male fertility (Barau et al., 2016; Zeng and Chen, 2019).

DNMT3A is expressed in two isoforms known as DNMT3A1 and DNMT3A2, both isoforms contain the PWWP and ADD domain which interact with H3K36me3, while only the latter interacts with unmethylated H3K4 (Dhayalan et al., 2010; Otani et al., 2009; Zeng and Chen, 2019). DNMT3A2 is the shorter isoform lacking the N-terminal before amino acid 223. However, while smaller it is the more predominant isoform that is responsible for *de novo* DNA methylation in embryonic stem cells (ESC) and was associated with euchromatin (Chen et al., 2002). In comparison, the DNMT3A1 is mostly expressed in differentiated cells and associates with heterochromatic regions and interacts with DNA through its N-terminal region (Chen and Chan, 2014; Suetake et al., 2011). Following this DNMT3B is highly isoformic with over 30 different variants, though some do not contain the catalytic domain, the inactive isoforms while associated with DNA methylation their mechanisms are still largely clouded. However, some studies have shown roles in modulating *de novo* methylation and their ability to act as an accessory protein in the absence of DNMT3L in somatic cells (Duymich et al., 2016; Gordon et al., 2013). The protein structure of these proteins is summarised in figure 1.1 below.

Figure 1.1

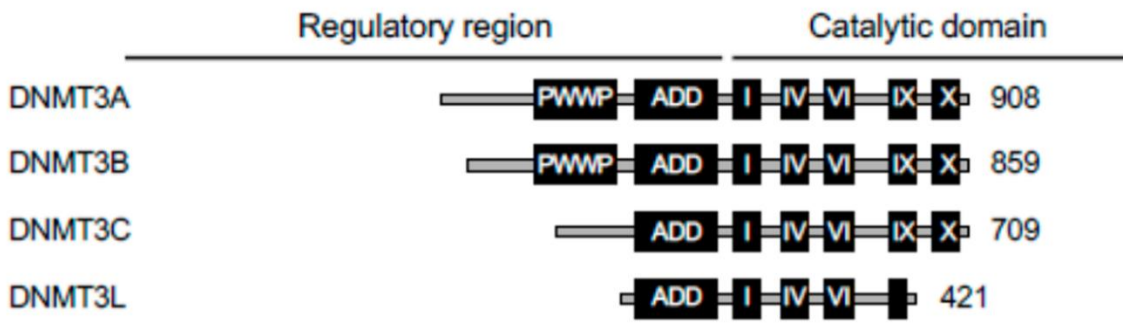


Figure 1.1: *De novo* methylation enzymes. Protein structure for the *de novo* DNA methyltransferases (DNMT) 3A, 3B and 3C. The DNMT3 enzyme family has overlapping catalytic domains and regulatory regions. The catalytic domain contains catalytic motifs (I-X) which are preserved in DNMT3A, 3B and 3C. Numbers represent the number of amino acids found in the mouse protein. PWWP, proline-tryptophan-tryptophan-proline domain; ADD, ATRX-DNMT3-DNMT3L domain. Edited from "DNA Methylation Reprogramming during Mammalian Development" "https://creativecommons.org/licenses/by/4.0/" (Zeng and Chen, 2019).

1.1.3 Maintenance DNA methylation enzymes

Methylation patterns are inherited during mitotic replication of DNA in somatic cells: DNMT1 is recruited to hemimethylated DNA where it maintains methylation (Jeltsch and Jurkowska, 2014; Wigler et al., 1981). In Dnmt1 KO mouse models it was observed that unlike Dnmt3a, Dnmt1 did not affect non-CpG methylation (Arand et al., 2012). The DNMT1 C-catalytic domain contains two sub-domains- the target-recognition domain (TRD) and the methyltransferase domain (Song et al., 2012). The TRD domain is important for recognising hemimethylated cytosines (Gujar et al., 2019). The N-terminal regulates protein interaction through multiple functional domains such as the replication foci targeting sequence (RFTS) domain (also known as replication foci domain) which localises DNMT1 to the replication fork, the CXXC zinc finger domain that recognises unmethylated CpG DNA and the paired bromo-adjacent homology (BAH) domain (Callebaut et al., 1999; Leonhardt et al., 1992; Pradhan et al., 2008; Zeng and Chen, 2019). DNMT1 contains an autoinhibitory loop which is an interposed stretch of acidic amino acids between the CXXC and BAH1 domain that is present at the DNMT1 active site (Song et al., 2011). This Auto-inhibitory loop along with the RFTS binding with the enzyme Ubiquitin Like With PHD And Ring Finger Domains 1 (UHRF1) that binds to hemimethylated DNA may explain the faithfulness of DNMT1 and hemimethylated DNA binding (Bostick et al., 2007). It has been suggested that the RFTS

domain blocks the active site of DNMT1, the RFTS is then displaced when DNMT1 forms the complex with UHRF1 and hemimethylated DNA (Edwards et al., 2017).

UHRF1 is a protein that is involved with the maintenance of DNA methylation, UHRF1 colocalises and interacts with DNMT1 at the replication fork (Sharif et al., 2007). The Ubiquitin Like (UBL) domain of UHRF1 was able to interact with the RFTS domain of DNMT1 allowing the two to bind, this binding was also seen to be essential for DNMT1 enzymatic activity (Li et al., 2018). As mentioned previously UHRF1 binds to hemimethylated DNA, it carries out this function through its Su(var)3-9, Enhancer of zeste and Trithorax (SET) and Really interesting new gene (RING) finger-associated (SRA) domain (Avvakumov et al., 2008). Furthering this the SRA domain is also known to interact with the RFTS domain of DNMT1 and increase its activity by 5-fold without the SRA DNA binding (Bashtrykov et al., 2014; Zeng and Chen, 2019). UHRF1 can recognise the histone marks H3R2 and H3K9me3/2 through the cooperation of the plant homeodomain (PHD) and tandem tudor domain (TTD) domains, with the PHD domain having a more dominant role (Arita et al., 2012). UHRF1 has an autoinhibitory mechanism in which a polybasic region (PBR) between the SRA and RING domains acts as a competitive inhibitor to H3K9me2/3 binding to the PHD domain, this is inhibition is prevented through the binding of phosphatidylinositol 5-phosphate binding to BRP resulting in a rearrangement of the domains (Gelato et al., 2014). The RING finger of UHRF1 is important for H3 ubiquitination due to its functionality as an E3 ubiquitin ligase (Citterio et al., 2004). The protein structure of DNMT1 and UHRF1 is summarised in figure 1.2 below.

Figure 1.2

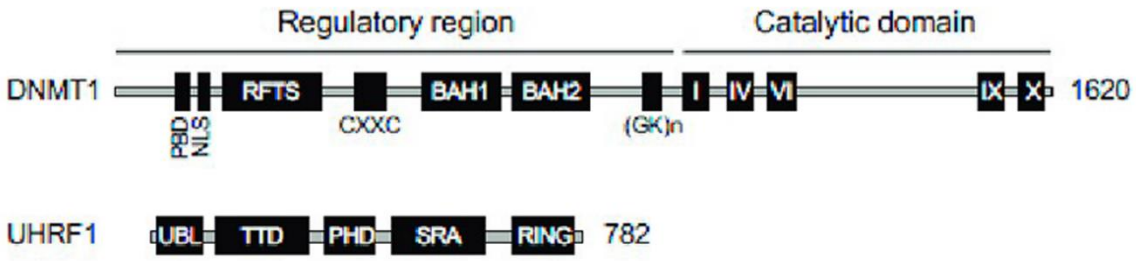


Figure 1.2: DNA maintenance methylation enzymes. Protein structure for the maintenance enzymes DNA methyltransferase 1 and ubiquitin−like with plant homeodomain and interesting new gene finger domains 1 (UHRF1). DNMT1 regulatory region and the catalytic domain is labelled. The catalytic domain contains catalytic motifs (I-X). Numbers represent the number of amino acids found in the mouse protein. PBD, proliferating cell nuclear antigen (PCNA)-binding domain; NLS, nuclear localization signal; RFTS, replication foci targeting sequence; CXXC, a cysteine-rich zinc finger domain; BAH, bromo-adjacent homology domain; (GK)n, glycine/lysine repeats; UBL, ubiquitin-like domain; TTD, tandem Tudor domain; PHD, plant homeodomain; SRA, Su(var)3-9, Enhancer of zeste and Trithorax (SET) and RING finger-associated; RING, really interesting new gene domain. Edited from "DNA Methylation Reprogramming during Mammalian Development" "https://creativecommons.org/licenses/by/4.0/" (Zeng and Chen, 2019).

1.1.4 Development of the early mouse embryo

In order to understand better the context in which DNA methylation changes occur during development, it seemed useful to provide a brief summary here of embryogenesis in the mouse. The first interval in mouse embryo development is preimplantation, this interval is where the inseminated oocyte implants to the uterine wall and establishes pregnancy. This interval lasts for 4.5 days containing a total of seven cleavage divisions, in which the late blastocyst consisting of three distinct cell lineages the trophectoderm (TE), inner cell mass (ICM) and primitive endoderm (PrE) is implanted on to the uterus wall (Mihajlović and Bruce, 2017). Upon fertilisation, the sperm is rapidly demethylated prior to the first cleavage while the oocyte methylation is gradually lost during division steps as described in Figure 1.3. The first two cleavage divisions lasts approximately 20 hours, which is considerably longer than the later divisions which only last around 12 hours (Artus and Cohen-Tannoudji, 2008). The zygote is not transcriptionally active until a small burst of activation by the end of the 1 cell-stage, this is then followed by a major activation upon completion of the 2 cell-stage this process is known as zygote genome activation (ZGA) (Schultz, 1993). Until ZGA the zygote relies on maternal mRNA, which is then mostly degraded by the end of the second cell-stage but some maternal mRNA will persist until the end of preimplantation (Johnson and

McConnell, 2004). A second major wave of activation occurs in the zygote at 4 cell-stage, this activation peaks at the 8 cell-stage and involves the translation of genes involved in morphological changes during late preimplantation development (Hamatani et al., 2004; Hamatani et al., 2006). It is around this time that demethylation of the oocyte is completed (Figure 1.3), there is also a final wave of transcript activation resulting in expression of genes involved in cavitation and cell-fate and lineage segregations resulting in the transition to blastocyst. The late blastocyst then leaves the zona pellucidae and implantation occurs at E4.5 (Wang and Dey, 2006).

At implantation (E4.5) the embryo has reestablished methylation levels through *de novo* synthesis which will then be maintained by Dnmt1 (Figure 1.3). The epiblast begins to proliferate into the blastocyst cavity due to the physical block imposed by the implantation site (Copp, 1979). Simultaneously the proamniotic cavity is being formed, the formation of the cavity is believed to be due to apoptotic signals from the visceral endoderm causing cell death (Coucouvanis and Martin, 1995). The PrE gives rise to the parietal endoderm and the visceral endoderm which play important roles in the development of the yolk sac as they make up the visceral and parietal yolk sacs components respectively of the mouse yolk sac (Xenopoulos et al., 2012). The development of the proamniotic cavity, parietal endoderm and the visceral endoderm and cell proliferation of the epiblast lead to the morphological changes observed in the pre-gastrula at E5.5 to E6.0 prior to gastrulation (Kojima et al., 2014). Gastrulation begins usually at E6.5 once the threshold number of epiblast cells have been reached, however, reaching the threshold before E6.5 earlier, as can be seen in double-sized embryos, does not initiate gastrulation any earlier, suggesting a chronological control. However, if the required number of epiblast cells has not been reached by E6.5 due to cell loss or disruption of proliferation, gastrulation is delayed until sufficient numbers of epiblast cells have been reached (Tam and Behringer, 1997).

Gastrulation (E6.5-E7.5) is the phase that results in specific morphologic transformation through cell proliferation and differentiation and morphogenetic movements. During the Gastrulation stage, primordial germ cells begin to be specified bringing about the full circle of genomic reprogramming with complete widespread demethylation across the genome and reestablishment of methylation including the imprints which are established as a lifelong

mark (Figure 1.3). Epithelial-to-mesenchymal transition is an important process for gastrulation in which epiblasts changes to mesenchymal cells at the primitive streak, resulting in the three primary germ layers (a transient embryonic structure made of epiblast cells that have started mesenchymal transition but still connected to the epiblast) (Nakaya et al., 2013). The epiblasts cells that undergo epithelial-to-mesenchymal transition at the proximal end of the embryo, become part of either the definitive endoderm germ layer or form a new tissue germ layer known as the mesoderm, while those that remain in the epiblast become the ectoderm with the mesoderm and endoderm contributing to many of the adult tissues (Acloque et al., 2009). Immediately after gastrulation begins early organogenesis (E8.0–E8.5): during this time the neural plate and heart tube begin to form (Cao et al., 2019a). In the next few days of organogenesis (E9.5–E13.5), the embryo begins to take its distinctive shape in which the head, limbs, heart and spinal cord can be identified (Mitiku and Baker, 2007). The remaining phase in embryo development following organogenesis is fetal growth and development and finally birth (Ko, 2001).

1.1.5 DNA demethylation

DNA demethylation is the process in which the 5-methylcytosine (5-mC) is converted to a cytosine, this can occur through two processes which can be passive and active demethylation. Passive demethylation occurs when DNA methylation is lost during the synthesis of a new DNA strand (Rougier et al., 1998). Active demethylation is when methylation is lost outside DNA synthesis, the exact mechanism is not fully elucidated, however, it is known that the ten-eleven translocation (TET) enzymes are involved (Bhutani et al., 2011; Paroush et al., 1990; Tahiliani et al., 2009). TET was found to have a hydroxylating activity to convert 5-mC to 5-hydroxymethylcytosine (5-hmC) (Tahiliani et al., 2009). Later the TET enzymes were found to be able to further convert 5-hmC to 5-formylcytosine (5-fC) then finally 5-carboxylcytosinen (5-caC) (Ito et al., 2011). The derivatives 5-fC and 5-caC are pathways for the thymine DNA glycosylase pathway (TDG) along with the BER pathway completes the demethylation of 5-mC to an unmethylated cytosine (Maiti and Drohat, 2011; Weber et al., 2016).

The TET enzymes carry out their hydroxylating activity by utilising the important co-factors Fe(II) and α-ketoglutarate (Wu and Zhang, 2014). The Tet enzymes have different expression levels through development to somatic tissues in mice. Tet3 is highly expressed in oocyte and preimplantation embryos until the blastocyst stage in which Tet1 and Tet2 show the highest expression. During differentiation to the germ layers, Tet2 and Tet3 expression is up-regulated or remains unchanged. Then finally in adult somatic tissue, Tet2 and to some extent Tet3 show high levels of expression (Rasmussen and Helin, 2016). Tet3 appears to play an irreplaceable role in embryonic development as the mouse KO model of Tet3 results in 100% embryonic lethality, while Tet1 and Tet 2 mouse KO models resulted in no developmental phenotype (Dawlaty et al., 2011; Gu et al., 2011; Quivoron et al., 2011). Interestingly the Double KO (DKO) of Tet1 and Tet2 is still tolerable in mice, though results in hypermethylation and compromises imprinting (Dawlaty et al., 2013). The same lab then carried out a triple KO (TKO) of the Tet enzymes in mESCs, this resulted in misregulation of genes, increased promoter methylation, loss of 5-hmC and impaired ESC differentiation further confirming the role of the Tet enzymes and 5-hmC in development and gene regulation (Dawlaty et al., 2014).

The oxidised modifications of 5-mC (5-hmC, 5-fC and 5-caC) have been increasingly investigated over the years and are important in many biological functions (Shi et al., 2017). 5-hmC is the most abundant oxidised derivative of 5-mC found at 3-5% of cytosine species in the human genome, while the more oxidised modifications are only found at 0.002–0.02% of all cytosine species (Dietzsch et al., 2018). Genomic regions in mESCs, 5-hmC and the more oxidised derivatives (5-fC and 5-cac) are found separately (Wu and Zhang, 2017). 5-hmC was discovered in 1952 shortly after the discovery of 5-mC (Wyatt and Cohen, 1952). Though 5-hmC is the result of the hydroxylating activity of the TET enzymes, it is important to note that increased expression of the TET enzymes does not correlate with the increased levels of 5-hmC suggesting that TET activity may be controlled by other factors (Szulwach et al., 2011). 5-hmC is found primarily at gene bodies of transcribed genes, however, 5-hmC can be variable across different tissues with similar gene expression suggesting that 5-hmC is dependent on tissue type (Nestor et al., 2012). Though the function of the gene body enrichment of 5-hmC is still to be determined (Lio and Rao, 2019), aberrant 5-hmC can be seen in tumorgenesis, stress response and neurological and psychiatric disorders (Chouliaras et al., 2013; Dong et al., 2012; Jin et al., 2011; Papale et al., 2017).

1.1.6 Methylation reprogramming

Methylation reprogramming occurs during gametogenesis and again throughout embryonic development, Widespread demethylation first occurs in the primordial germ cells (PGCs) during gametogenesis, this is the first drop of methylation as seen in Figure 1.3B. During this time the diploid PGC's undergoes demethylation removing all previous methylation marks including imprinting (Hajkova et al., 2002). Before the primordial germ cells can become sperm or oocyte, dependent on the sex, methylation is then reestablished which includes the imprinting marks known as gametic DMRs (gDMRs) (Figure 1.3), most of which are maternally methylated and are then protected from any further reprogramming and retained as a life-long mark (Guibert et al., 2012; Lees-Murdock et al., 2003; Proudhon et al., 2012). However, an exception to this is transient DMRs which are protected during gametogenesis though are susceptible to reprogramming during embryonic development leading to tissue-specific gDMRs (Proudhon et al., 2012; Rutledge et al., 2014). The next reprogramming occurs during fertilisation: almost immediately after fertilisation occurs, as observed by the different degree of gender-specific methylation as seen in Figure 1.3B, the sperm is rapidly demethylated before the first round of division occurs, while the oocyte methylation occurs passively over several divisions (Oswald et al., 2000). As mentioned previously the gDMRs are protected from demethylation during embryonic development, this is due to the binding of two Krüppel-associated box containing zinc finger proteins (ZFP), ZFP57 and ZFP445, though other effectors may also be involved (Ferguson-Smith, 2011; Takahashi et al., 2019). ZFP57 and ZNF445, when bound to a gDMR by targeting UHRF1, maintain methylation and H3K9me3 histone modifications (Strogantsev et al., 2015; Takahashi et al., 2019). Repetitive DNA and a subset of single-copy genes also avoid DNA demethylation during embryonic development (Borgel et al., 2010; Guibert et al., 2012; Lane et al., 2003; Lees-Murdock et al., 2003). After birth, the establishment of further DMRs can be established, these are known as somatic DMRs (sDMR).

Figure 1.3

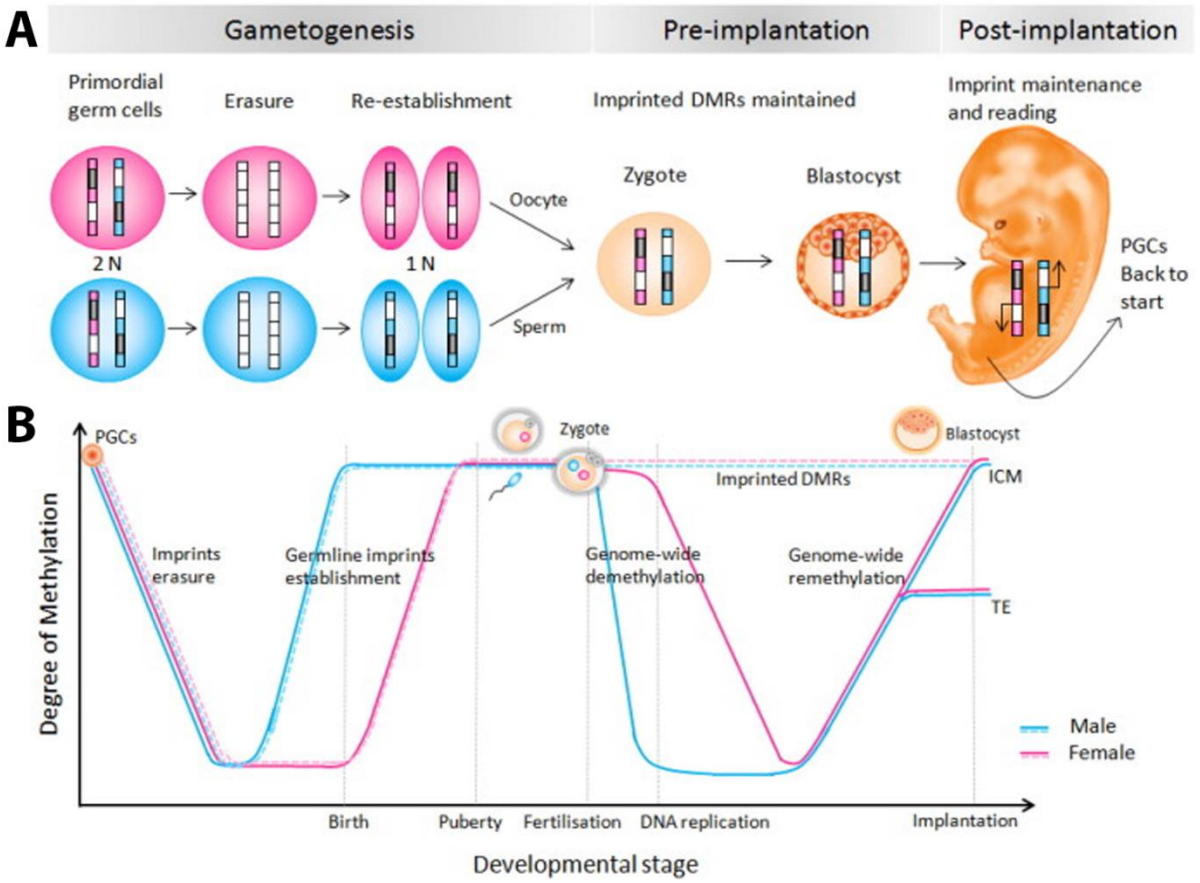


Figure 1.3: DNA methylation reprogramming throughout development. (A) Methylation reprogramming at gametic differentially methylated regions (gDMR) throughout development. Primordial germ cells (PGC) (Diploid) contain the methylation imprints as inherited from the parents. As PGCs undergo gametogenesis methylation gDMR marks are lost. As the PGC matures towards their respective germ cell depending on gender they become haploid and the methylation marks are established. After fertilisation, the zygote is formed and begins embryonic development, as it progresses through the pre-implantation phase gDMR methylation is maintained. Progressing through embryonic development, imprint gDMRs continue to be maintained and even transcribed and a monoallelic fashion. PGC formation and gametogenesis occurs during post-implantation beginning the reprogramming loop from the gametogenesis stage. Black arrows represent transcription from the allele. (B) DNA methylation reprogramming genome-wide and imprints during embryonic development in mice. The timescale in the graph overlap with the events from figure 1.3A. At the beginning of gametogenesis in PGCs widespread methylation erasure occurs, however, as the PGCs mature towards their respective germ cell methylation is re-established this occurs at birth for sperm and later in life for oocyte (Puberty). Upon fertilisation another genome-wide (Solid) DNA methylation erasure event occurs genome-wide, however, the imprinted gDMRs (dashed) are protected. As the embryo develops through pre-implantation methylation is gradually re-established genome-wide. Upon implantation, the inner cell mass (ICM) is known to carry more methylation than the trophectoderm (TE). Reprinted from "The role of imprinted genes in humans" "https://creativecommons.org/licenses/by/3.0/" (Ishida and Moore, 2013).

1.1.7 Environmental factors and DNA methylation

DNA methylation is a very stable epigenetic regulator, though it should be noted that it is susceptible to changes through environmental and genetic factors (Hannon et al., 2018). The effect of environmental factors varies depending on age, sex and tissue (van Dongen et al., 2016). It was noticed that the ageing process leads to overall hypomethylation in the genome (Heyn et al., 2012). With this in mind, studies then looked further into methylation as a predictor of age (Bocklandt et al., 2011). The Horvath clock predicts the age of an individual by the methylation pattern depending on the tissue which leads to the concept of DNA methylation age (DNAmAge) (Horvath, 2013). Several studies went on to show that increase of DNAmAge compared to actual age resulted in a term denoted as accelerated ageing (AA) which results in a higher risk of mortality or disease, with an increase of 1-year DNAmAge there was an increased risk of cancer (Jylhävä et al., 2017; Zheng et al., 2016a).

DNAmAge can be affected by environmental factors. One environmental factor that has been shown to cause AA is diet. A study investigating the association of socio-economic impact with epigenetic differences found that those on the lower end of the socioeconomic scale were seen to have methylation associated with AA that they contributed to the higher consumption of red meat (McClelland et al., 2016). Smoking is a further factor known to cause hypomethylation, these marks of methylation from smoking persist even after an individual quits smoking (Shenker et al., 2013). Studies have even shown that the effects on methylation can be seen in several tissues and can even be seen in newborns related to maternal smoking during pregnancy (Joubert et al., 2012). However, interestingly though smoking has been shown to have profound effects on methylation and age-related disease the literature is conflicting if there is a correlation between AA and smoking (Beach et al., 2015; Gao et al., 2016; Marioni et al., 2015).

Training and exercise have been observed to have beneficial effects that can be correlated with changes of methylation at promoters and enhancers of loci related to the health-enhancing phenotype (Lindholm et al., 2014). Further to this, it was seen that folic acid supplementation during the second and third trimester resulted in changes of methylation related to brain development in cord blood, suggesting that maternal folate status may play

an important role in neurodevelopment in offspring (Caffrey et al., 2018). Furthering this study the authors showed that in the same cohort the children in which their mothers took folic acid supplementation had higher emotional intelligence and resilience (Henry et al., 2018).

1.2 Functional roles for DNA methylation

DNA methylation plays many essential roles in mammalian development, this includes inactivation of the X chromosome, silencing of repetitive element activity, silencing of the germ-line genes and cell differentiation. Ageing and the direct environment are well known to correlate with changes in methylation. Methylation plays an important role in the normal regulation of several gene classes, which I will outline in this section; likewise, aberrant methylation can lead to various diseases such as cancer and imprinting disorders.

1.2.1 Inactivation of the X chromosome

In mammals, there are two sex chromosomes X and Y, with males having XY and females containing XX. The X chromosome contains ~1100 genes, whereas the Y contains ~100 genes, this would lead to a genetic imbalance in Females (Heard and Disteche, 2006; Pinheiro and Heard, 2017). To compensate for the extra X chromosome, females inactivate the X chromosome to prevent transcription, in a process known as X inactivation (Lyon, 1961). X inactivation occurs during ESC differentiation after implantation. X inactivation is regulated by the X inactivation centre which contains the lncRNA known as X inactive-specific transcript (Xist) (Brown et al., 1991). Xist is expressed from the inactive X chromosome where it coats the chromosome leading to the subsequent enrichment of repressive histone marks and methylation (Heard *et al.*, 2001). The Xist lncRNA appears to be regulated by methylation as it was shown that in Dnmt1 KO mESCs and embryos that methylation at the promoter of the active X chromosome is required for stable repression (Panning and Jaenisch, 1996). It should be noted that the selection of the chromosome to be inactivated appears to be completely random (Takagi et al., 1982). In addition to regulating Xist, DNA methylation is important for repressing CGI-containing genes on the X^i (Gendrel et al., 2012; Grant et al., 1992; Lock et al., 1986). Despite the repressive epigenetic modifications some genes appear to 'escape' x-

inactivation. Some genes escape x-inactivation: studies vary but 8-15% genes on the inactive x chromosome are reported as consistently expressed while another 10-32% of genes are variable depending on the individual and tissue. It is currently unclear if the escaped genes avoid the silencing signal or are reactivated after silencing (Posynick and Brown, 2019). In comparison x-inactivation in mouse has fewer genes escaping, ranging from 3-7% depending on the strain and tissue (Balaton et al., 2015). Genes that escape x-inactivation are found throughout the x chromosome however, they are most prevalent at regions homologous between sex chromosome, known as the pseudoautosomal regions (PAR) (Berletch et al., 2011). The escaped genes have been associated with the disease phenotype in a female monosomy disorder known as Turner's syndrome. The PAR genes *CSF2RA* and *SHOX,* in particular*,* have been suggested to be involved in the placental defects and short stature disease phenotypes respectively (Clement-Jones, 2000; Urbach and Benvenisty, 2009).

1.2.2 Silencing of 'selfish' DNA

There are many types of repetitive elements in the human genome which makes up 66-69% of the human genome (de Koning et al., 2011). The 'selfish' DNA term is used for transposable elements, of which there are two types. Class 1 also known as Retrotransposons makes up 42% of the human genome and rely on their mRNA being reverse transcribed and then reincorporated randomly back into the genome (Finnegan, 2012; Lander et al., 2001). Class 2 also known as DNA transposons use the 'cut and paste' mechanism which involves the DNA transposon being cut out and reincorporated randomly back into the DNA sequence (Smit and Riggs, 1996). Transposable elements can have detrimental effects by causing loss of gene activity (Hancks and Kazazian 2012). Methylation plays an important role in maintaining transposon elements in a repressed state and it has even been speculated that the evolution of methylation was primarily to repress the transposable elements (Yoder et al., 1997). Loss of methylation has been shown to derepress long interspersed nuclear element-1 (LINE 1) in human ESCs using a DNMT inhibitor 5'Aza-dC as well as in mice deficient in DNMT1, respectively (Walsh et al., 1998; Woodcock et al., 1997). Though transposable elements are usually detrimental they also play a role in genomic evolution as a source of genetic variation, as some genes contain transposon-like elements that are important for correct regulation and function of the gene (Jordan et al., 2003; van de Lagemaat et al., 2003). This source of genetic

variation caused by transposable elements has also been associated with the establishment of several imprinted genes in humans such as *RTL1* (Grothaus et al., 2016). Transposition events usually result in inverted or truncated copies of the transposable elements that are inactive and accumulate in the genome.

LINE1 is an active and the most abundant non–long terminal repeat retrotransposable element in mammals making up 20% of the genome in human and mouse (Sassaman et al., 1997; Yang and Wang, 2016). Studies suggest that there are up to 3000 intact copies of *LINE1* in mouse and up to 100 copies in human (DeBerardinis et al., 1998; Mandal and Kazazian, 2008). *LINE1* retrotransposable activity is carried out through its two open reading frames (ORF1 and ORF2). ORF1 encodes nucleic acid chaperone activity while ORF2 is required for reverse transcriptase and endonuclease activity (Feng et al., 1996; Martin et al., 2005; Mathias et al., 1991). *LINE1* also contains a 5' and 3' untranslated region (UTR), the 3' UTR is punctuated by a poly(A) tail, while the 5' end functions as an internal promoter containing a CpG island which is usually methylated to silence expression (Beck et al., 2011).

Human endogenous retroviruses (HERVs) are a type of Long terminal repeat (LTR) retrotransposon which are remnants of ancient retroviruses that are largely immobile and inactive and are inherited in a Mendelian fashion (Jern and Coffin, 2008). While most HERVs are inactive, there remains a small number of "young" HERV which are almost intact and can be transcribed nearly in full. Additionally, despite most HERVs loss of retrotransposon activity and function, the LTR may still contain functional regulatory regions which can then act as alternative promoters for nearby genes (Cohen et al., 2009). Thus, repression of both full-length and truncated HERV by DNA methylation is crucial to the organism. DNA transposons only make up 3% of all elements and are highly diverse with 120 families which are categorised into 5 superfamilies, however, these DNA transposable elements are now mostly inactive remnants (Feschotte and Pritham, 2007; Plasterk et al., 1999).

1.2.3 Germline silencing

Genome-wide studies relatively recently uncovered an additional subset of genes that were heavily methylated at their promoters in somatic tissue leading to their subsequent repression, while in the germline these promoters were hypomethylated and subsequently expressed (Meissner et al., 2008; Weber et al., 2005). During embryonic development repression of the germline genes is one of the most robust and coordinated methylation mechanisms (Smith and Meissner, 2013). Dnmt3b is found at the promoter sites of the germ-line specific genes, where it establishes methylation during ESC differentiation at implantation in mouse embryos at E6.5 (Borgel et al., 2010). DNMT3B is recruited by the transcription repressor E2F6 to the promoter site of the germ-line specific genes (Velasco et al., 2010). Leseva *et al* (2013) showed for a group of germ-line specific genes that the absence of E2f6 resulted in aberrant reactivation. The overexpression of Dnmt3b and subsequent hypermethylation at the promoter site of *Stag3* was not sufficient to silence *Stag3* expression either in E2f6 deficient mouse embryonic fibroblasts (Leseva et al., 2013). DNMT1 plays an important role in maintaining the methylation at the germline promoters, as Dnmt1-null mouse somatic cells failed to methylate the promoter regions which subsequently resulted in the upregulation of germline genes in somatic tissues (Maatouk et al., 2006). Failure to repress germ-line genes in somatic tissue can have detrimental effects, though the expression is normal during the germ-line, the abnormal expression has been identified in many different tumour types (Simpson et al., 2005).

1.2.4 Genomic Imprinting

Genomic imprinting was first discovered in 1960 in Sciara but was referred to as a controlling element in sex chromosome behaviour, it wasn't until the early to mid-1980s that imprinted genes were suspected to be present in the mammalian genome (Crouse, 1960; McGrath and Solter, 1983; Surani and Barton, 1983; Surani et al., 1984). It has been well established that imprinted genes play an essential role in embryo and placental biology. They also play an important role in congenital function and metabolism, with aberrant regulation of Imprinted genes leading to imprinting disorders, congenital conditions and an increased risk of cancer (Peters, 2014). There have been over 100 imprinted genes identified in mammals (Luedi et al., 2007; Nikaido et al., 2003). Imprinted genes are usually found in well-conserved clusters

and predominantly expressed in the placenta. Paternal expressed imprinted genes usually promote growth while maternal expressed imprinted genes usually inhibit growth (Renfree et al., 2013). How Imprinted genes came to be is not fully understood but there are two major theories in literature which include the predominant Kinship theory and the maternal-offspring coadaptation theory (O'Brien and Wolf, 2017). However, the evolution of imprinting is not within the scope of this thesis.

Imprinting DMRs can be established in both the germline and after fertilisation in somatic tissues, which are known as germline DMRs (gDMR) or somatic DMRs (sDMR) respectively as described in Figure 1.4 (Pervjakova et al., 2016). The first observation of an sDMR was at the promoter region of the mouse *Igf2r* which was seen to have paternal methylation outside the germline in mouse (Stöger et al., 1993). A further DMR that was observed to be established outside the germline is located at the maternally expressed Cdkn1c gene in mouse in which the promoter region was also paternally methylated (Bhogal et al., 2004). The best-described model for an ICR is the regulation of the *H19/Igf2* genes in mouse. The ICR at the *H19* gene is usually methylated on the maternal allele and unmethylated on the paternal allele leading to maternal expression of H19 and paternal expression of the Igf2. A paternally methylated sDMR was seen to be important for further silencing of the paternal expression of the H19 gene, shown in Figure 1.5 (Bartolomei et al., 1993; Ferguson-Smith et al., 1993). However, it should be noted that the same group went on to show that deletion of the gDMR does not result in the loss of methylation at the sDMR suggesting that once established the sDMR is stable and that the gDMR is only required for its establishment (Srivastava et al., 2003).

Figure 1.4

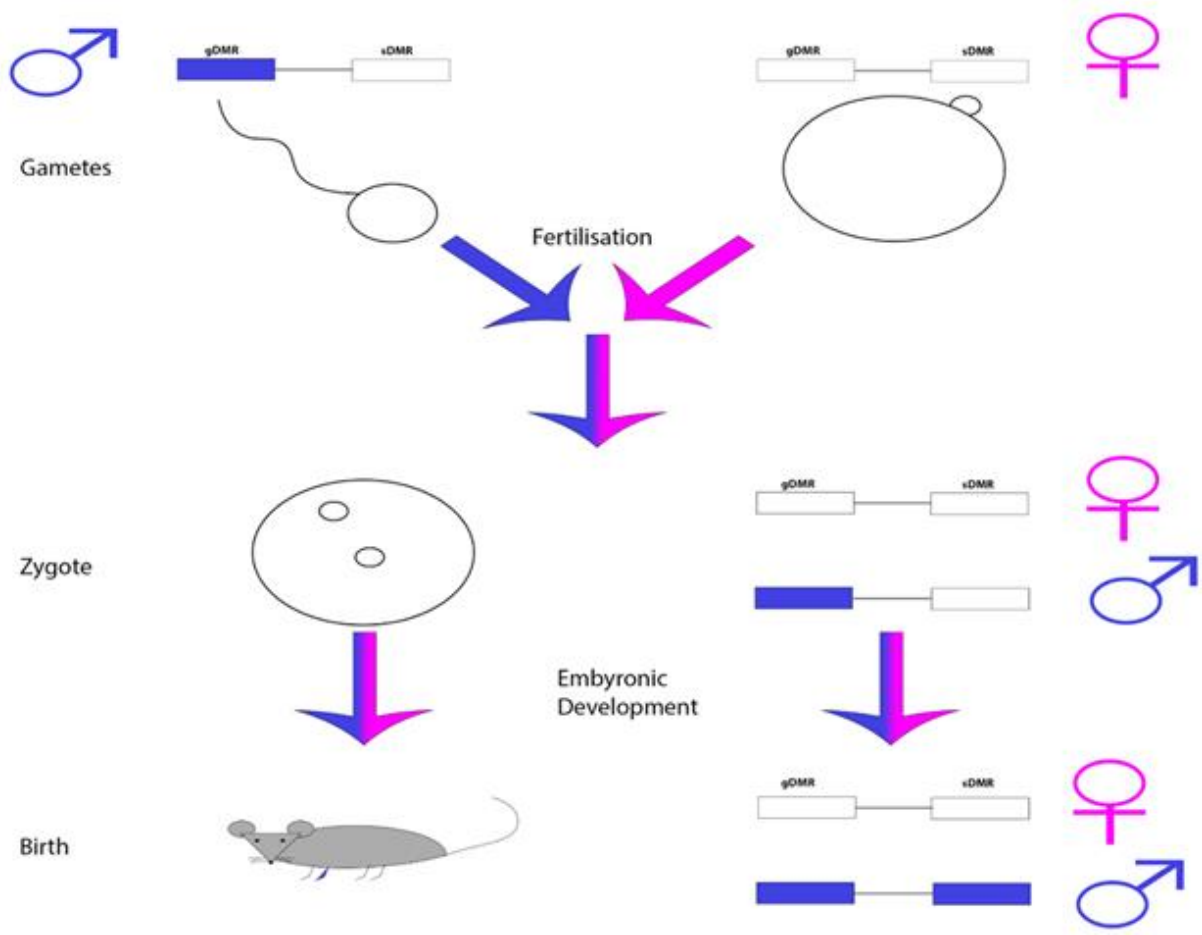


Figure 1.4: Schematic diagram describing the establishment of imprinting gDMRs and sDMRs in mice. During gametogenesis DNA methylation undergoes reprogramming, it is here that the gametic DMRs (gDMR) are established and maintained as a lifelong mark. On the formation of a zygote, methylation undergoes a second round of reprogramming: however, the imprinted gDMRs are protected during this phase. It is the embryonic stage of methylation reprogramming that can lead to the establishment of the somatic DMR (sDMR).

Figure 1.5

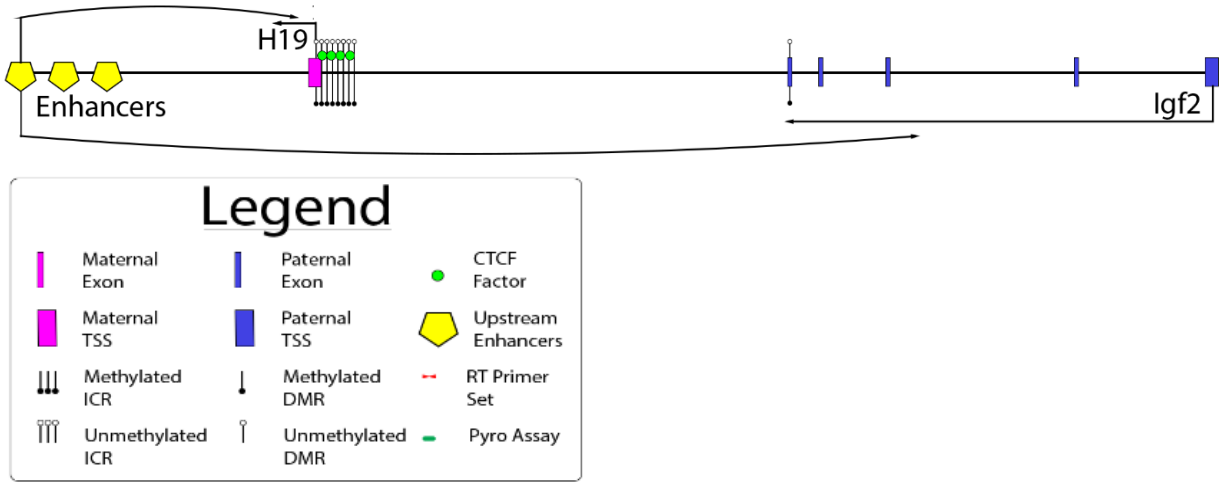


Figure 1.5: Schematic diagram describing the ICR and sDMR functionality at the H19/Igf2. On the Maternal Allele, the ICR and sDMR are unmethylated, this opens the CTCF binding site, allowing CTCF to bind and prevent the upstream enhancers from interacting with *Igf2*. Therefore, the enhancers instead interact with *H19* resulting in the expression of H19 and repression of *Igf2*. On the paternal allele, the ICR and sDMR are methylated, this blocks the CTCF binding site preventing CTCF binding. This allows the enhancers to interact with *Igf2*, leading to its subsequent expression, while the sDMR is important for the silencing of *H19*. Legend describes all symbols found within the figure.

The environment plays an important role in altering methylation and has been even known to alter the methylation at Imprinted genes. One environmental factor that can affect Imprinted genes is maternal nutrition, as maternal nutrition is critical for foetal development it is widely explored across many systems (Kappil et al., 2015). Maternal folate status has shown to have an inverse relationship with DNA methylation at a tested set of imprinting DMRs (Hoyo et al., 2014). Further to this, it was seen that maternal smoking during pregnancy and paternal drinking would affect the methylation at imprinting DMRs such as *MEG3* in the offspring and *H19* in the sperm respectively (Markunas et al., 2014; Stouder et al., 2011). Assisted reproduction technologies (ART) have also been seen to have a large impact on genomic imprinting. A population study in Japan showed an association between ART and imprinting disorders such as BWS and SRS, further to this widespread disruption of the imprinting gDMRs was also observed after the use of ART (Hiura et al., 2012). However, Deangelis, Martini and Owen (2018) have suggested that the risk of imprinting disorders in ART is low and screening is unwarranted due to studies being unable to differentiate if the risk of imprinting disorders is due to the fertility treatment or rather the fertility issues in the sample population.

1.2.5 Differentiation

ESCs and PGCs are the building blocks to essentially all cell types, during differentiation remarkable changes in morphology and function can be seen usually largely determined by distinct gene expression. This distinct expression usually involves the silencing of self-renewal specific genes and the expression of tissue-specific genes (Hackett et al., 2012). *De novo* DNMTs establish the tissue-specific methylation patterns that lead to cell differentiation during embryonic implantation (Wu and Sun, 2006). Further to this double knockout (DKO) mouse ESCs are unable to undergo differentiation *in vitro*, further confirming the importance of methylation in tissue-specific stem cells and/or differentiated cells (Okano et al., 1999). It is possible to induce pluripotency in differentiated cells to give induced pluripotent stem cells (iPSCs) which closely resemble the epigenetics of normal ESCs including DNA methylation after continuous passaging (Nishino et al., 2011). The imprinting status of most imprints is conserved through induced pluripotency, though if imprinting is lost it can not be fixed through continuous passage (Hiura et al., 2013).

Reprogramming of somatic tissue to iSPCs has a large impact with benefits for both clinicians and researchers while minimising ethical conflicts. However, the efficiency of reprogramming is low (0.1-3%) with a high financial cost (Gomes et al., 2017). This low efficiency can be partly attributed to the epigenetic memory of the source cells, that is partially explained by incomplete DNA promoter methylation (Kim et al., 2010; Ohi et al., 2011). Supporting this Mikkelsen *et al.*, (2008) showed that treatment with the DNMT 1 inhibitor 5-aza-2'-deoxycytidine (AZA) resulted in higher cell death at early treatment stages, but a four-fold increase of embryonic-stem-cell like colonies at later stages which led the author to suggest that demethylation of 1 or more loci is important for late-stage reprogramming. A further method to increase the efficiency of cell reprogramming is treatment with vitamin C: by alleviating cell senescence, promoting H3K36me2/3 demethylation through Jhdm1a/1b and preventing loss of imprinting at the Dlk1/Dio3 loci through blocking Dnmt3a binding (Esteban et al., 2010; Stadtfeld et al., 2012; Wang et al., 2011). Interestingly loss of function studies of Dnmt3a and 3b showed that *de novo* methylation is dispensable when inducing pluripotency (Pawlak and Jaenisch, 2011).

1.3 DNA methylation changes associated with Disease

1.3.1 DNA methylation and disease Functionality

As discussed previously correct DNA methylation regulation is important as it may lead to disorders due to misregulation of genes required for normal bodily function. As mentioned previously deviations from the normal DNAmAge can lead to diseases such as cancer, diabetes, cardiovascular disease and neurological disorders (Fransquet et al., 2019). In this section, the epigenetic changes in a number of diseases will be discussed.

Cancer initiation, promotion and progression is the result of global genetic and epigenetic abnormalities (Hanahan and Robert A, 2017). DNA methylation is often associated with regulating the processes contributing to the neoplastic transformation such as Genome-wide hypomethylation resulting in activation of retrotransposons and cancer-testis antigen (CTA) genes (Jasek et al., 2019). Retrotransposal elements and CTA genes have been shown to have the ability to stimulate oncogenic pathways involved with angiogenesis, cell proliferation, metastasis and immortality as well as inhibiting tumour suppressor pathways involved with apoptosis, controlling growth inhibition signals and genome integrity (Van Tongelen et al., 2017). A further study by Greißel *et al.*, (2015) showed loss of methylation at the *LINE-1* elements in high-grade carotid artery stenosis. *LINE-1* hypomethylation was also observed in a study of 292 myocardial infarction cases and a longitudinal analysis of ischaemic heart disease and stroke showed hypomethylation at the *LINE-1* elements was associated with higher risk and subsequent mortality, further supporting hypomethylation at the involvement of the *LINE-1* element in CVD (Baccarelli et al., 2010; Guarrera et al., 2015).

Epigenetic alterations have been seen in metabolic disorders such as type 2 diabetes (T2DM) (Bansal and Pinney, 2017). Methylation was first associated with diabetes in candidate gene studies which showed patients with T2DM had increased methylation at the promoters of the peroxisome proliferator-activated receptor γ and peroxisome proliferator-activated receptor-gamma coactivator-1 alpha genes (Barrès et al., 2009; Ling et al., 2008). Interestingly loss of imprinting at the DMR on chromosome 11 known as KvDMR (also known as DMR2), may

mediate susceptibility to diabetes through the aberrant expression of CDKN1C, KCNQ1 and KCNQ1OT1 during development (Travers et al., 2013).

1.3.2 Imprinting Disorders

Imprinting disorders, as mentioned previously occur due to aberrant regulation of imprinted genes. Imprinting disorders are characterized by overlapping clinical features such as abnormal development, growth and metabolism usually caused by molecular disturbances affecting the imprinted genes. A list of imprinting disorders and their respective affected gene can be seen in Table 1.1. Imprinting disorders can be caused by four main molecular changes which include uniparental disomy (UPD), genetic mutations at the imprinted loci, chromosomal rearrangements and epimutation. However, the phenotypic changes that occur are usually dependent on the affected allele (Eggermann et al., 2015b). UPD is when both chromosomes are inherited from the same parent and causes an imbalance of imprinted genes on the affected chromosome. UPD is known to be associated with almost every known imprinted disorder (Eggermann et al., 2015a). Genetic mutations are known to be the cause of a few imprinting disorders, the Imprinted gene *CDKN1C* when mutated is known to cause the imprinting disorders Silver-Russell syndrome (SRS) and Beckwith-Wiedemann syndrome (BWS) (Brioude et al., 2013; Hatada et al., 1996). Chromosomal rearrangements such as deletions, duplications and translocations can alter regions important for the regulation of imprinted genes or the imprinted gene itself, these rearrangements would then cause the aberrant expression of the respective gene and therefore lead to an imprinting disorder related to the imprinted gene affected. One example is the deletion of the ICR at the 11p15.5 region which is often referred to as ICR 2, the deletion of this ICR on the maternal chromosome is associated with BWS (De Crescenzo et al., 2013). Epimutations account for 50% of cases for imprinting disorders, this occurs when methylation at the imprinting DMRs are abnormal (Eggermann et al., 2015b).

Loss of imprinting has been seen to cause increased tumorigenesis in mice, in humans the imprinting region *H19/IGF2* has been known to be aberrantly expressed in several cancers (Lim and Maher, 2010). A further Imprinted gene *RTL1* which is present at the *DLK1* imprinting domain has been seen to be overexpressed in patients with hepatocarcinogenesis (Riordan

et al., 2013). As well as cancer, imprinted genes also have been linked to psychiatric disorders such as autism spectrum disorders (Fradin et al., 2010). In mice the paternally expressed *Grb10*, when deficient in the brain, alters social behaviour in adults (Garfield et al., 2011).

Table 1.1

Imprinting Disorder	OMIM	Common mutations	Chromosome	Parental Inheritance	Clinical Features	
Beckwith-Wiedemann (Wang et al., 2020)	*130650*	*IGF2/H19* DMR GOM KvDMR LOM UPD(11)pat	11p15	Paternal	• abdominal wall defects • visceromegaly • hyperinsulinemic hypoglycemia	• Pre- and postnatal overgrowth • Macroglossia • pinna abnormalities
Silver-Russell syndrome (Spiteri et al., 2017)	180860	*IGF2/H19* DMR LOM UPD(11)mat UPD(7)mat	11p15 Chr 7	Maternal	• clinodactyly • relative macrocephaly • ear anomalies • skeletal asymmetry	• Pre- and postnatal growth restriction • short stature • triangular face
upd(20)mat Syndrome (Hjortshøj et al., 2020)	N/A	UPD(20)mat	Chr 20	Maternal	• Pre- and postnatal growth restriction	• feeding difficulties
Transient Neonatal Diabetes Mellitus (Touati et al., 2019)	601410	*PLAGL1* DMR LOM UPD(6)pat	6q24	Paternal	• renal abnormalities • developmental delay • congenital heart diseases • hypothyroidism	• Prenatal growth restriction • transient diabetes • hyperglycemia without ketoacidosis • umbilical hernia
Prader-Willi syndrome (Touati et al., 2019)	176270	UPD(15)mat *SNURF/SNRPN* DMR GOM	15q11	Maternal	• intellectual disability • Behavioural problems • risk of developing psychosis • Hypogonadism	• Pre- and postnatal growth restriction • hypotonia • feeding difficulties • hyperphagia
Angelman syndrome (Maranga et al., 2020)	105830	UPD(15)pat *SNURF/SNRPN* DMR LOM	15q11	Paternal	• neurological problems • hyperactivity • attention deficit	• Intellectual disability • Microcephaly
Kagami-Ogata syndrome (Ogata and Kagami, 2016)	608149	UPD(14)pat *IG-DMR* GOM *MEG3* DMR GOM	14q32	Paternal	• increased coat-hanger angle • omphalocele • placentomegaly • developmental delay	• Prenatal overgrowth • protruding philtrum • polyhydramnios • feeding difficulties

Pseudohypoparathyroidism Ib (Tafaj and Jüppner, 2017)	103580	UPD(20)pat *GNAS-NESP* DMR LOM *GNAS-XL:Ex1* DMR LOM *GNASA/B* DMR LOM	20q13	Paternal	• postnatal growth restriction • rounded face • brachydactyly • ectopic ossifications • Intellectual disability	• renal resistance to Parathyroid hormone • absence of any features of Albright's hereditary osteodystrophy • mild resistance to Thyroid-stimulating hormone
Precocious puberty syndrome (Abreu et al., 2013; Dauber et al., 2017)	615346	*MKRN3* mutations *DLK1* mutations	15q11 14q32		• Precocious puberty in girls younger than 9	• Precocious puberty in boys younger than 8
Temple syndrome (Prasasya et al., 2020; Temple et al., 1991)	616222	UPD(14)mat *IG-DMR* LOM	14q32	Maternal	• muscular hypotonia • precocious puberty • feeding difficulties	• Pre- and postnatal growth restriction • small hands and feet
Birk-Barel Syndrome (Barel et al., 2008)	612292	*KCNK9* mutations	8q24		• feeding problems • dolichocephaly • short philtrum	• congenital hypotonia • variable cleft palate • delayed development
Schaaf-Yang syndrome (Schaaf et al., 2013)	615547	*MAGEL2* mutations	15q11		• infantile feeding problems • distal joint contractures • gastroesophageal reflux • chronic constipation	• intellectual disability • developmental delay • autism spectrum disorder • neonatal hypotonia • skeletal abnormalities

Table 1.1: Overview of known imprinted disorders clinical features and molecular abnormalities. Uniparental disomy (UPD) is shown which chromosome is affected by number in brackets and whether is maternal or paternal by mat/pat respectively. Recent reviews or relevant studies when reviews were not available are cited below each imprinting disorder. DMR, differentially methylated region; GOM, gain of methylation; LOM, Loss of methylation; pat, paternal; mat maternal; Chr, chromosome.

1.3.3 Mutations in DNA methyltransferase genes and disease

As shown throughout this chapter the regulation of the methylome is essential for normal biological function throughout the body, therefore it is not surprising that mutations in any of the three DNMT's are associated with disease states.

A syndrome called immunodeficiency, centromere instability and facial anomalies-1 (ICF-1) (OMIM 242860) is caused by loss-of-function mutations at DNMT3B (Xu et al., 1999). The mutations are found mostly located within the catalytic domain leading to abnormal DNA binding, SAM utilisation, SAM binding and homo-oligomerization (Moarefi and Chédin, 2011). Missense and splice site mutations do occur in the N-terminal in some ICF patients, leading to an inactive truncated protein (Jiang et al., 2005). ICF is an extremely rare autosomal disease that was first described in 1978, with only around 50 cases reported worldwide up to 30 years later (Ehrlich et al., 2006; Hulten, 1978). ICF patients suffer from recurrent severe infections usually resulting in death before adulthood due to agammaglobulinemia caused by peripheral terminal B-cell blockage (Blanco-Betancourt et al., 2004). ICF most typical feature is chromosome instability at chromosomes 1, 9 and 16 due to gains and losses of chromosome arms due to loss of methylation at classic satellites 2 and 3. This loss of methylation is observed in all cells, however, the chromosome instability is only present in specific cell types (Edwards et al., 2017). Further loss of methylation is also observed at a subset of the germline genes promoters *MAEL* and *SYCE1*, CpG islands on the inactive x-chromosome, SST1 (a family of microsatellites), subtelomeric regions and *D4Z4* DNA repeats (Velasco and Francastel, 2019). Velasco *et al.*, (2014) suggest that the hypomethylation of *MAEL* and *SYCE1* and subsequent derepression of these germ-line genes may act as specific and robust biomarkers for the diagnosis of ICF.

DNMT1 mutations are associated with two neurodegenerative adult-onset disorders, Hereditary Sensory Neuropathy with dementia and hearing loss type 1E (HSAN1E) (OMIM 614116) 10 and Autosomal Dominant Cerebellar Ataxia, Deafness and Narcolepsy (ADCA-DN) (OMIM 604121) (Klein et al., 2011; Winkelmann et al., 2012). HSAN1E to date has been associated with 9 mutations located frequently within exon 20 while only 4 mutations have been associated with ADCA-DN and only exclusively found in exon 21, both exons encode for

parts of the RFTS domain (Norvil et al., 2019). With both diseases being linked to mutations in the RFTS domain of DNMT1, it is not surprising that they have many overlapping clinical features, such as cognitive decline by late ’40s, sensorineural hearing loss, mild to severe neuropathy and similar age-onset of disease and life expectancy. One exception is the prominence of narcolepsy associated with ADCA-DN that is only found in some cases of HSAN1E (Baets et al., 2015). A study utilising whole-genome bisulphite sequencing, showed genome-wide hypomethylation, at most CpG islands, promoters, exons, L1, L2, Alu and satellite repeats and simple repeat sequences. Imprinted genes were most affected with more hypomethylated CpGs than in other genes. However, some regions did appear to be hypermethylated (Sun et al., 2014). Kernohan *et al.*, (2016) showed a similar trend using the 450k array in ADCA-DN patients and suggested that the DNMT1 mutations cause similar methylation profiles in both disorders with hypomethylation at repetitive elements and site-specific hypermethylation at specific gene promoters and CpG islands.

DNMT3A mutations are associated with a number of disorders such as germ-line mutations in Tatton-Brown-Rahman syndrome (TBRS) (OMIM 615879) and microcephalic dwarfism (MD), while somatic mutations are observed in Acute Myeloid Leukemia (AML) (OMIM 601626) ((Heyn et al., 2019; Ley et al., 2010; Lin et al., 2018; Tatton-Brown et al., 2014). TBRS is an overgrowth syndrome that arises due to mutations in the germline with 40 different variants found in any of the three functional domains. TBRS presents with facial abnormalities, tall stature and intellectual disability (Norvil et al., 2019). Genome-wide profiling of 15 TBRS patients observed accelerated epigenetic ageing with hypomethylation at genes associated with development, differentiation, morphogenesis and malignancy predisposition pathways, this was seen to be most severe in the patient with the R882H mutation that is seen in >20% of AML cases (Jeffries et al., 2019). R882H mutations result in 80% reduction of DNMT3A activity in the mutant enzyme, which is also known to inhibit the WT from forming tetramers which represent the most active form of DNMT3A (Russler-Germain et al., 2014). AML patients suffering from mutations in R882H in the MTase domain are seen to have genome-wide hypomethylation at specific CpGs predominantly at homeobox genes (Qu et al., 2014). The R882H mutations in AML patients have also been observed to be predictive of a poorer prognosis, with AML R882H patients having an inferior response to aclarubicin treatment (Yuan et al., 2019). While the R882H mutation is the most

common occurring DNMT3A mutation in AML, it should be noted that 15-20% of DNMT3A mutations in AML cases are outside the R882H codon, but rather they contain insertions, truncations or single-copy deletions mutations, however, these mutations are less investigated in the literature (Bruno et al., 2019). A recent study by Heyn *et al.*, (2019) described two PWWP gain of function germ-line mutations W330R and D333N in microcephalic dwarfism. Using the Infinium® MethylationEPIC BeadChip (EPIC) array hypermethylation was observed at Polycomb-regulated regions associated with developmental genes. The increase of methylation was found in conjunction with the depletion of H3K27me3 and H3K4me3. This led the authors to believe that the PWWP mutation impaired the Polycomb repressive complex 2 (PRC2) binding (Heyn et al., 2019).

1.3.4 The Infinium HumanMethylation BeadChip Array and RNA-seq

The use of genome-wide technology has helped researchers to gain a greater understanding into epigenetic regulators, such as RNA-seq, Methylation BeadChip array (450k and EPIC), ever since the human genome project, the technological advances within genomics has exploded, accelerating a change in biology and medical research fields. With the introduction of the 1000 genome project and HapMap, researchers were provided with an extensive catalogue of human variation, opening the door to more in-depth investigations with genome-wide association studies (GWAS) (Altshuler et al., 2010; Auton et al., 2015). Epigenome-wide association studies (EWAS) have also flourished over recent years, in particular following the introduction of the Illumina Infinium methyl-seq (reduced representation/whole genome bisulphite sequencing), histone/chromatin enrichment with and transcriptome analysis (RNA-seq) to name a few (Chaitankar et al., 2016; Sandoval et al., 2011). Utilising the genomewide technologies mentioned above-allowed researchers to understand epigenetic changes from environmental effects to disorders (Barletta et al., 1997; Lindholm et al., 2014). This has been especially progressive for imprinting disorders leading to the discovery of novel DMRs and imprinted genes as well as candidate imprinted regions and genes involved in imprinting disorders (Court et al., 2013; Prickett et al., 2015).

1.4 Summary and Hypothesis

Genome-wide methylation NGS utilises sodium bisulphite treatment of DNA, this treatment converts unmethylated cytosines to thymine which is ultimately converted to uracil, which can be then quantified (Altshuler et al., 2010). Illumina Infinium methylation BeadChip arrays utilise a probe-based fluorescent system that hybridises to bisulphite treated DNA, resulting in a signal. There are two types of probes: Type I probes use a paired system -one binds to the methylated locus and the other unmethylated locus; while type II only contains 1 probe which binds to both the methylated and unmethylated locus (Zhou et al., 2017). Illumina Infinium Methylation BeadChip arrays have a high level of reproducibility and reliability with probes covering RefSeq genes, promoter CpG islands and covers all known imprinting clusters in both the HumanMethyation450 (450K) and MethylationEPIC (850K) arrays (Carmona et al., 2017). RNA-seq is a type of high throughput sequencing technology (sequences multiple DNA molecules in parallel), it utilises cDNA synthesised directly from an RNA sample to provide the transcriptional status at a specific time at when it was extracted. RNA-seq is carried out in 3 stages, library preparation to manipulate the RNA to be compatible with sequencing, sequencing and finally analysing with bioinformatic tools (Qian et al., 2014).

It is clear that methylation plays a major role in genomic imprinting and genomic imprinting like disorders. However, there is still much left unknown about the mechanism of how these regions and their related genes are regulated. Imprinting disorders have a unique mechanism such as parental specific methylation controlling monoallelic expression, as well as being protected from the genomewide demethylation during reprogramming of the zygote. Understanding the imprinted genes and regions fully will allow for more accurate diagnosis and may help pave the way for effective treatments for imprinting disorder patients. This thesis will aim to expand our understanding behind genomic imprinted and retrotransposable regions. Firstly, we hypothesise that the 450k array can be used to characterize and ‘score’ genomic imprinted genes to differentiate between normal and abnormal levels of methylation to identify DMR disturbances. To answer this I will utilise publicly available datasets and analyse them using R studio and use galaxy to create a methylation index for normal methylation variability of the genomic imprinted regions and test against datasets for patients diagnosed with genomic imprinting disorders. Our second hypothesis is that the

widespread loss of methylation caused by the depletion of DNMT1 in the hTERT-1604 cell line will disrupt the regulation of the imprinted regions. To test this we aim to investigate the 450k array datasets generated by the Walsh lab using genome coordinates for imprinted DMRs found in the literature and confirm the changes of methylation using pyrosequencing technology, qPCR will then be used to confirm any consequent functional changes. The third and final hypothesis is that the point mutation in the PHD domain of Uhrf1 will result in the abrogation of the H3K binding and subsequent aberrant epigenetic regulation of the genomic retrotransposable and imprinted elements. We aim to investigate the retrotransposable and imprinted elements through the use of wet-lab experiments designed to analyse the methylation and transcriptional status of these elements in PHD mutant homozygotic embryos at E8.75 and E9.5.

Chapter 2. General Materials and Methods

2.1 Tissue Culture

The adherent fibroblast hTERT immortalised human lung fibroblast-1604 (hTERT-1604) (Ouellette, 2000) and the adherent HT29 and HCT116 epithelial colon cancer cell lines was cultured in high glucose 4.5g DMEM media (Gibco, Paisley, UK) supplemented with 10% fetal bovine serum (Gibco, Paisley, UK) and 2x NEAA (Gibco, Paisley, UK). Previous members of the CPW laboratory generated the respective DNMT1 and UHRF1 stable knockdowns in hTERT-1604s using a pSilencer construct (Ambion, Huntingdon, UK) with integrated DNMT1/UHRF1 targeting shRNA. For selection of shRNA cells, media was supplemented with 150µg/ml hygromycin B (Invitrogen, Paisley, UK) and culture was carried out the same as wild type hTERT-1604s; hygromycin B treatment was halted at least 48hours before any experimental procedure. Conditions for all cell culture was 37°C, 5% CO_2 in a humidified incubator. Downstream analysis was carried out on cells at passage #22-24.

2.2 Cell passaging

Cell passaging was carried out at ~70-80% confluency when media was removed and x2 washes was carried out with PBS (Oxoid Ltd). Cells in a 90mm tissue culture plate (Iwaki) were deattached with 1ml 1x trypsin-EDTA (Gibco, Paisley, UK) incubation at 37^{o}C; time of treatment required for disassociation of the was reliant on the cell line in use. FBS supplemented media was used to end treatment. Cells were pelleted in a universal tube by centrifuging for 5 minutes at 1200rpm; the supernatant was then removed and the pellet was gently agitated to unclump cells. The pellet was resuspended in 5ml of complete media; a haemocytometer counting chamber (Neubauer) was utilised to count the cells. The volume of cells used for passaging was formulated based on the number of cells counted to ensure culturing consistency; media was added to a total of 10ml media. Additional components (5-Azacytidine and 5-aza-2'-deoxycytidine) were added as required before re-incubation.

2.3 Preparation of cell stocks

Cells were frozen down when ~70-80% confluent. Trypsanisation and cell counting was carried out as explained previously, 5 cell stocks could be created from cells at 70-80% confluency in a 90mm plate. The volume required for $5x10^6$ cells was centrifuged at 1200rpm for 5min in a universal tube to form a pellet; the supernatant was removed and the pellet was gently agitated to unclump cells then resuspended in 2.5ml of media. 2.5ml ice-cold freezing solution (80% FBS; Invitrogen, Paisley, UK, 20% DMSO; Sigma-Aldrich, Dorset, UK) was added dropwise to the cell/media mix whilst agitating the cell solution to ensure even distribution of the freezing solution. The cell solution was then split into 5x1ml individual cryovials (Davidson & Hardy Ltd, Belfast, UK) and immediately, transferred into a Mr Frosty™ Freezing Container (ThermoFisher, Loughborough, UK) and stored in a -20^{o}C freezer overnight. The cryovials were placed at -80^{o}C for at least 24hr and retained at -80^{o}C for short term storage or transferred to the liquid nitrogen facility within the University's cryostorage for long term storage.

2.4 DNA extraction and Bisulphite conversion

Cell pellets were resuspended in 500µl of lysis buffer [50 mM Tris pH 8, 0.1 M EDTA (both Sigma-Aldrich) incubated overnight with rotation at 55°C, 0.5% SDS, 0.2 mg/ml proteinase K (Roche, West Sussex, UK)]. The phenol/chloroform/isoamyl alcohol (25:24:1 pH8, Sigma-Aldrich) extraction method was utilised to isolate the DNA. Once extracted UV absorbance measurements at 260/280 and 260/230 nm using a Nanodrop UV spectrophotometer were used to quantify and examine the quality of DNA (Labtech International, Ringmer, UK). 250/500ng of DNA depending on the quantity of DNA available was used for bisulphite conversion with the EZ DNA Methylation Kit (Zymo, Cambridge, UK) according to the manufacturer's instructions.

2.5 Bisulfite PCR

Amplification was carried out using the PyroMark PCR kit (Qiagen, Crawley, UK) with 2 µl bisulphite-converted DNA, 12.5 µl MasterMix, 2.5 µl CoralLoad Concentrate, 1.25 µl each primer (10 µM) and 5.5 µl nuclease-free water (Qiagen, Crawley, UK) using the following conditions: 15 min at 95 °C followed by 45 cycles of 94 °C for 30 s, 56 °C for 30 s, 72 °C for 30 s and a final elongation step of 72 °C for 10 min. Primers were designed in-house using the PyroMark Assay Design software 2.0 before synthesis (Metabion, Germany) some assays were ordered pre-designed from Qiagen (Crawley, UK) were as described.

2.6 Agarose gel electrophoresis

Gel electrophoresis is used to determine the success of the PCR by the presence of a band at the expected size. 1µl of the produce was mixed with 5µl of loading dye (20% glycerol, 0.05% orange G; both from Sigma-Aldrich, Dorset, UK) before being separated through a 1% agarose gel by electrophoresis. 1% Gel prepared by heating 100ul 1x TAE buffer with 1g agarose electrophoresis grade powder (Apollo Scientific, Cheshire, UK) and 2ul GelRed Nucleic Acid Gel Stain, 10,000X in DMSO (Cambridge Bioscience, Cambridge, UK) until clear then left to set in a cast. 1x TAE buffer was made using a 50x concentrated stock which consisted of 2M Tris-base, 0.05M EDTA (Sigma-Aldrich, Dorset, UK), 1M acetic acid (VWR, Lutterworth, UK). To estimate the size of the product, samples were run alongside the 100bp ladder (Invitrogen, Paisley, U.K). Agarose gels were imaged by UV transillumination on the Bioview UV light (Biostep, Jahnsdorf, Germany).

2.7 Pyrosequencing

A pyro PCR product was generated as seen in the bisulphite PCR above, however, a primer contained a biotin label which is dependent on the read direction of the assay. The pyro PCR was then analysed using Pyromark™ Q24 pyrosequencer (Qiagen, Crawley, UK). To the PCR product, 2µl beads (GE Healthcare, Chalfon St. Giles, UK), 40µl binding buffer (Qiagen, Crawley, UK) and 18µl nuclease-free water (Qiagen, Crawley, UK) was added. Samples were placed on a shaker for 5min at 1400rpm allowing for attachment of the sepharose beads to the pyro PCR product. The sepharose bead-bound PCR products were washed with 70%

ethanol (5 sec), denaturation solution (5 sec) (0.2M NaOH; Sigma-Aldrich, Dorset, UK) and 1x wash buffer (10 sec) (Qiagen, Crawley, UK), the PCR products were then released onto a PyroMark™ Q24 plate (Qiagen, Crawley, UK) which contained 0.3µM sequencing primer in 25µl annealing buffer (Qiagen, Crawley, UK) at the pyromark workstation. The pyromark was placed on a heat block at 80°C for 2min and left to cool at RT for >5min, allowing the sequencing primer to anneal to the pyro PCR product. The pyromark Q24 cartridge was then loaded with enzyme and substrate solutions and nucleotides from PyroMark Gold Q24 Reagents (Qiagen, Crawley, UK) following the manufacturers' instructions. Both the samples and cartridge was placed into the Pyromark™ Q24 pyrosequencer (Qiagen, Crawley, UK) and samples were sequenced. The results were viewed using the associated PyroMark™ Q24 software, v.2.0.6 (Qiagen, Crawley, UK).

2.8 Pyrosequencing Statistical analysis

The mean methylation across all CpG sites analysed in the assay was calculated for each CpG assay. The standard deviation (SD) was calculated and the standard error of the mean (SEM) was calculated using the SD. In some assays the methylation value for each of the CpG sites able to be assessed by the assay was used for the SD and SEM; this is indicated in the figure legend when this was the case. Significance was carried out using the non-parametric Mann–Whitney U test.

2.9 RNA extraction, cDNA synthesis and RT-qPCR

RNA was extracted using the RNeasy kit (Qiagen, Crawley, UK), according to the manufacturer's instructions. cDNA synthesis was carried out with 250/500ng RNA used in combination with 0.5 µg random primers (Roche, West Sussex, UK), 40 U RNaseOUT 0.5 µM dNTPs (Invitrogen, Paisley, UK) 1× RT Buffer (Fermentas, Cambridge, UK) and RevertAid reverse transcriptase (Fermentas, Cambridge, UK) made up to a final volume of 20 µl using RNase-free water (Qiagen, Crawley, UK). Reactions were carried out in a thermocycler with conditions—25 °C for 10 min, 42 °C for 60 min and 70 °C for 10 min. One microlitre cDNA per well on a 96-well plate (Roche) was used for RT-qPCR with SYBR Green reagent and the remaining cDNA was stored at −80 °C. RT-qPCRs was performed using a LightCycler 480 Instrument II (Roche, West Sussex, UK). Gene expression was normalised to HPRT or GAPDH

for both human and mouse studies as indicated in the figure legend. Relative expression was calculated by the ΔΔCT method [40]. Each RT-qPCR contained 1× buffer, 0.4 mM dNTPs, 50 µM primers, 0.01 U Taq DNA polymerase (Invitrogen, Paisley, UK) and nuclease-free water (Qiagen, Crawley, UK). The general thermocycler conditions are as follows—94 °C for 3 min, followed by 30 cycles of 94 °C for 30 s, 63 °C for 1 min, 72 °C for 1 min with a final elongation step of 72 °C for 4 min.

2.10 qPCR Statistical analysis

Gene expression was significance was determined by two-tailed, unpaired t-test function on excel. SD was calculated using the mean fold changes in >1 experiment which in turn was used to determine SEM. In some experiments only 1 triplicate value was used for the SD and SEM; this is indicated in the figure legend when this was the case.

2.11 R processing

Gene Expression Omnibus (GEO) dataset files were downloaded and processed using the RnBeads (Assenov et al., 2014) methylation analysis package (v1.0.0) through the GSE dataset identifier. In some cases, iDATs were downloaded manually for the GEO datasets and then processed with RnBeads in the same manner. Processing of the data included quality control, filtering out of probes containing SNPs and checking for hybridisation performance. The data was then normalised using the SWAN method in minfi (Aryee et al., 2014) after background subtraction with methylumi.noob. The data was then extrapolated from RnBeads as a bed file with methylation as absolute β levels and uploaded to USEGalaxy where it was subjected to the in-house workflow and subsequent analysis in SPSS.

2.12 UseGalaxy Workflow

The online bioinformatic software usegalaxy.org allows for the creation of online workflow tools. The Walsh lab designed a workflow for extracting the data at a list of genes or locations from processed beadchip array datasets uploaded as bed files containing absolute β levels of methylation. This workflow extracts the methylation value at each probe at the region or gene specified, the workflow will then calculate the probe number, mean, std, median, min and

max which is then used to work out the standard error of mean and significance. In some figures, the individual probes were taken and statistical analysis was carried out using SPSS using the non-parametric Mann–Whitney U test.

Chapter 3. Determination of the potential utility of Methylation BeadChip arrays in the clinical diagnosis of imprinting disorders

3.1 Introduction

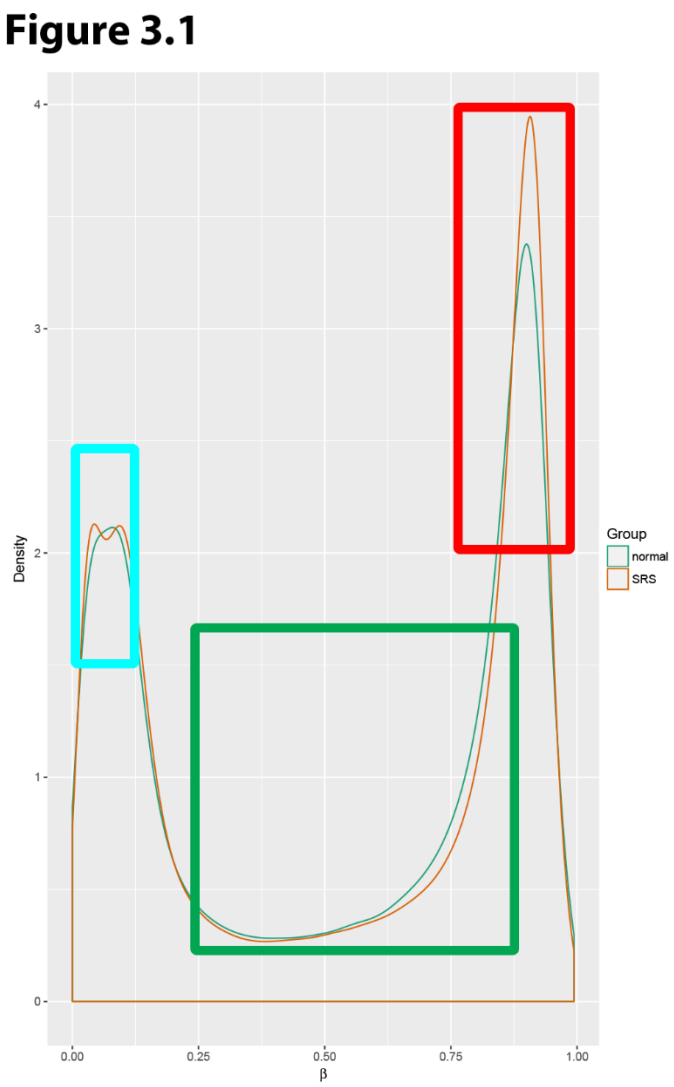
3.1.1 Clinical Diagnosis of imprinting disorders

As established throughout this thesis, imprinting disorders are a group of congenital disorders defined by molecular aberrations at imprinted genes/regions. The molecular disturbances observed in imprinting disorders are not the only feature shared among these disorders, many imprinting disorders also share similar clinical features involved in growth, metabolism and development. Due to the broad phenotype spectrum among imprinting disorders and the overlap of clinical features can result in patients being mis- and/or underdiagnosed due to a lack of efficient diagnostics and clinical management (Eggermann et al., 2015c). When imprinting disorders are suspected, the first test carried out would usually be a DNA methylation analysis of the ICR (Paulsen and Ferguson-Smith, 2001). This is due to the methylation disturbances at the ICR due to single-locus imprinting disturbances (SILD), copy number variance (CNV), or UPD. DMRs are regions in which allele-specific methylation occurs, a quantitative analysis of DNA methylation at imprinted DMRs carried out by Woodfine, Huddleston, and Murrell 2011 suggested DMRs established in the germ-line were quite stable and would normally fall within a range of 35-65% in healthy individuals. It is also important to consider the methylation differences between gDMRs and sDMRs, as gDMRs typically display very high levels of DNA methylation on the methylated allele (90–100%) and very low levels of DNA methylation on the unmethylated allele (0–10%). In contrast, significantly more variations in DNA methylation patterns are observed at secondary DMRs (Nechin et al., 2019).

Unlike the more common methylation control elements observed in the genome which are normally hypermethylated or hypomethylated as described in the density plot in Figure 3.1. Figure 3. 1 is a bimodal graph with the probe density on the Y-axis and beta methylation value across the bottom. The first peak represents the number of hypomethylated probes, while the second peak represents the number of hypermethylated probes in both the healthy and disease datasets. The area of the graph between the two peaks represents the intermediately

methylated probes. As imprinted DMRs are intermediately methylated it is no surprise that the SRS patients display a decrease of intermediately methylated probes and an increase of hypermethylated probes due to their disturbances at a subset of DMRs.

Figure 3.1: Density plot comparing the methylation distribution across probes from the Infinium 450k beadchip array in blood samples from SRS patients and healthy controls. In normal conditions density plots will present with the peaks showing here in which highly methylated probes are largest in density (Red), followed by lowly methylated probes (Blue), then a low density of intermediate probes (Green). Imprinted regions would fall under intermediately methylated probes due to having methylation levels around 50%. Therefore, it is understandable that the disruption of the imprinted regions caused by Silver-Russell Syndrome (SRS) patients would result in a decrease of intermediately methylated probes and an increase in highly methylated probes. Generated by Gareth Pollin (2020) using the dataset GSE55491 from Prickett et al (2015).



Methylation-specific PCR and southern blots were frequently used for the diagnosis of imprinting disorders. However, due to the labour intensity of the southern blot and qualitative-only data limitation of both methods, they have over time been replaced with methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) (Algar, 2019). MS-MLPA is carried out by the hybridisation of two gene-specific adjacent probes to the sample DNA, it utilises methylation-sensitive enzymes to digest the DNA along with DNA ligases for ligation, the product is then amplified using an automated sequencer as seen in

Figure 3.2. This is advantageous over the other methods as it is simple and several tests can be run simultaneously. Subsequent results are 1) quantitative; 2) return data which includes methylation at either one, two, or none of the alleles depending on the disturbance and finally 3) can show CNV, CNV analysis would have been previously ran separately along with the methylation-specific PCR and southern blots (Nygren et al., 2005). In patients with UPD, diagnosis is usually carried out by genotyping of microsatellite markers at the chromosome suspected to be affected in comparison to the two parental chromosomes, also single nucleotide polymorphism (SNP) arrays may be used as a diagnostic tool for UPD (Altug-Teber et al., 2005; Chen et al., 2019; Wilson et al., 2008). Diagnostic tests for imprinting disorders usually focus on one locus/region, with the two commonest being chromosomes 11 and 15 which I will talk about first. With the increase in studies using genome-wide technologies, however, it is being observed that many imprinting disorders are not restricted to the disease-specific regions but rather additional genetic and epigenetic disturbances can be seen at other imprinted loci and DMRs throughout the genome (Eggermann et al., 2014) and I will discuss these before looking at current diagnostics.

Figure 2

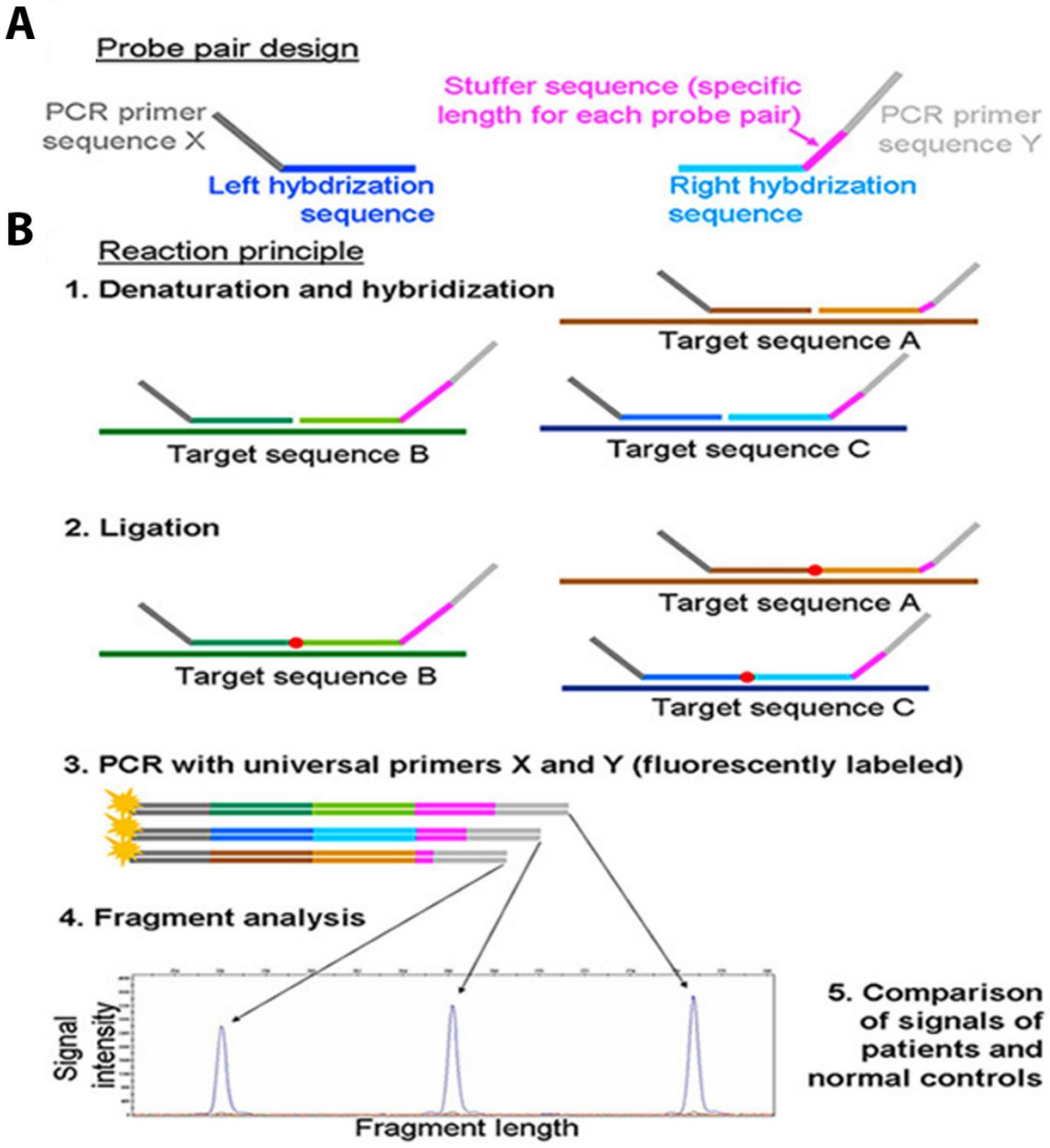


Figure 3.2: Experimental protocol for methylation-specific multiplex ligation-dependent probe amplification. (A) Each probe is designed with a target sequence for a restriction site of a methylation-sensitive endonuclease, a nonhybridizing tail containing and universal primer sequences (X and Y). A "stuffer sequence" is added to either probe to modulate the length. **(B)** The left and right probes will hybridize adjacently to the target sequence, after this ligation with the methylation-sensitive endonuclease is carried out but only cutting probes that have bound to unmethylated strands, leaving the methylated bound probes intact. The intact probes are then amplified using fluorescently labelled primers complementary to the primer sequences on the probes. Fragments are usually analysed by measuring signals from both digested and undigested samples, comparison of these two samples it is possible to determine the DNA methylation levels. reprinted from "DNA methylation analysis" "Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)" (Von Känel and Huber, 2013).

3.1.2 Chromosome 11 Imprinting disorders

The short arm on chromosome 11 (11p15) contains two very distinct imprinted clusters controlled by their respective ICRs. These two ICR's show opposite parental methylation and play an irreplaceable role in growth and development. The H19 maternally methylated DMR (also known as ICR1) controls the expression of the *H19*/*IGF2* loci and some nearby genes, the mechanism in how it controls expression at this region is explained in the General Introduction chapter. The *KvDMR* (also known as ICR2, KvLQT1-A andLIT1) is the contrasting paternal DMR, controls a cluster of genes around the *KCNQ1* and *CDKN1C* loci as well as nearby genes (Begemann et al., 2012). The *KvDMR* is not well understood in contrast to the *H19* DMR, though recent studies have shown a strong requirement for maternal transcription of the *KNCQ1* gene. Disruption of transcription of the *KNCQ1* gene has been shown to correlate with a failure to establish methylation at the *KvDMR* (Beygo et al., 2019a; Valente et al., 2019). Aberrant expression of the maternal or paternal genes at these ICRs can lead to maternal Silver-Russell Syndrome (SRS) and paternal Beckwith Wiedemann syndrome (BWS) imprinting disorders respectively (Õunap, 2016).

SRS is a Pre- and postnatal growth restriction disorder with many clinical features, it was first described by Silver *et al.*, (1953) and later by Russell (1954). Azzi *et al* (2015) suggested the most recent clinical scoring system in which they list 6 of the most common SRS phenotypes: if the patients match 4+ of the items then they likely have SRS, +2 SRS is suspected and less than 2 they do not have SRS with a sensitivity of 98%. This list includes (1) prenatal growth retardation; (2) postnatal growth retardation; (3) relative macrocephaly at birth; (4) protruding; (5) body asymmetry; (6) feeding difficulties. The SRS phenotype has been observed together with disturbances at different loci throughout the genome, with approximately 40% of cases occurring due to hypomethylation at the *H19* DMR with a subset of these patients suffering from MILD (7-9%) (Eggermann et al., 2016). The other primary cause of SRS cases is due to mUPD 7 (10%). Rare cases of SRS has also been observed in some patients with mUPD of chromosome 16 and 20, partial deletions on chromosomes 1, 11, 12 and 15 were also observed in rare cases of SRS (Azzi et al., 2015; Fokstuen and Kotzot, 2014). Due to the complex nature of this disease, 30-40% of patients are diagnosed based on clinical diagnosis with no underlying molecular defect found (Õunap, 2016).

BWS is a heterogeneous disease compared to SRS and is its polar opposite in terms of growth and molecular disturbances (Galerneau, 2018). BWS is primarily an overgrowth syndrome but linked with many other clinical features, it was first described in 1963 and 1964 by Beckwith and Wiedemann respectively (Beckwith, 1963; Wiedemann, 1964). A study on a set of Chinese BWS patients observed that 48.1% of cases were linked to loss of methylation at the maternally methylated *KvDMR*, while 11.1% of patients were associated with hypermethylation at the paternally methylated H19 DMR, 33.3% of patients suffered from upd(11)pat and finally, 7.5% of patients were seen with *CDKN1C* mutations (Luk, 2017). Diagnosis of BWS is made through the help of a clinical scoring system, a recent consensus statement defined a criterion using cardinal and suggestive features awarding 2 and 1 points respectively. Patients with >4 points are sufficient to be diagnosed with BWS while patients with >2 are recommended to have genetic testing and further investigation, those with <2 do not require genetic testing (Brioude et al., 2018). Unlike SRS only 13-15% of BWS patients are diagnosed based on clinical symptoms with no molecular disturbance (Õunap, 2016).

3.1.3 Chromosome 15 Imprinting Disorders

The PWS/AS region on chromosome 15 (15q11–13) is controlled by an imprinting centre (IC) pair which includes the PWS-IC and AS-IC, both the ICRs control a number of imprinted genes most of which are paternally expressed (Chamberlain and Lalande, 2010a). The PWS-IC contains a DMR and is found at the promoter region of the SNURF/SNRPN loci, the AS-IC does not contain a DMR and is found 35kb downstream from the PWS-IC. Though not imprinting control regions, the *MKRN3* and *NDN* promoter regions contain maternally methylated DMRs (Horsthemke and Wagstaff, 2008; Matsubara et al., 2019). There is strong evidence that the AS-IC drives maternal methylation at the PWS/AS region, recently it was reported that the AS-IC overlaps with oocyte-specific transcription start sites in which the transcripts involved may play an important role in maternal-specific methylation at PWS/AS region (Beygo et al., 2020; Lewis et al., 2015; Matsubara et al., 2019). Molecular defects at the PWS/AS imprinting region are observed in patients with PWS and AS, in which aberrant expression of the imprinted genes is seen (Rabinovitz et al., 2012). PWS was first described in 1956 by Andrea Prader it presents with clinical features such as hypotonia, hypogonadism, obesity and central nervous system dysfunction (Prader, 1956). The current criteria that is used as a guideline to

recommend patients for genetic testing, is split into different age groups, gradually increasing the number or severity of the common features involved with PWS. Genetic testing is suggested from birth to 2 years old if hypotonia with poor suck is observed while the criteria for 13 years of age through to adulthood includes common features like intellectual disabilities, excessive eating and hypothalamic hypogonadism and/or typical behaviour problems. Further criteria exist for 2-6 and 6-12-year-olds (Gunay-Aygun et al., 2001). The molecular defects in PWS is much more defined than the previously discussed imprinting disorders. About 70% of cases occur due to a regional paternal chromosomal deletion at 15q11-q13, while roughly 25% of cases are caused by maternal disomy of chromosome 15 and the remaining individuals have either epimutations or microdeletions at the PWS-IC (Butler et al., 2016). Interestingly the molecular defects do not cause different combinations or addition/removal of clinical features, though patients with 15q11-q13 maternal duplications appear to have more severe behaviour and psychological problems (Butler et al., 2004).

AS is a rare neurogenetic disorder first described by Harry Angelman in 1965, similar to PWS it presents with intellectual disability, though no other clinical features overlap (Angelman, 1965). The diagnosis of AM involves a criteria list of clinical features leading to suspicion, however, it can only be effectively diagnosed with molecular testing (Williams et al., 1995; Williams et al., 2006). Similarly, to PWS about 75% of AM patients contain a regional chromosomal deletion at 15q11-q13 but instead on the maternal chromosome. Interestingly parental UPD is associated with a small percentage (1-2%) of AM patients (Buiting et al., 2016). A further 1-3% can be caused by imprinting defects, either aberrant methylation or loss of the AS-IC, leading to loss of maternal imprinting. 10% of patients are associated with mutations at *UBE3A,* the majority of which are premature stop codons and the remaining patients are observed to have no identifiable defect (Beygo et al., 2019b). Only two paternally imprinted genes are present at the PWS/AS imprinted region, which are *UBE3A* and *ATP10A*: while UBE3A plays a role in cell cycle regulation, the function of ATP10A is currently not yet fully understood (Ehrhart et al., 2018).

3.1.4 Multilocus imprinting disorder

How MLID (Multilocus imprinting disturbances) occurs has not yet been established, however, it has been observed to be dependent on the tissue and the imprinting disorder. MLID has currently been investigated in syndromes such as BWS, SRS, transient neonatal diabetes mellitus (TNDM1) and Pseudohypoparathyroidism Ib (PHP1b) (Eggermann et al., 2015b). MLID was first observed by Mackay et al (2008) in TNDM patients with mutations within the ZFP57 gene resulting in variable but stable mosaicism of hypomethylation. The TNDM patients with the ZFP57 mutation present with classical TNDM as well as other, heterogeneous features, most notably developmental delay and congenital heart disease. The authors observed the first instance in which a global heritable global imprinting disorder that was compatible with life. A follow-up study by Boonen *et al.*, (2013) further investigated the implication of TNDM1 patients with a mutation in *ZFP57*, the authors observed consistent loss of methylation at the *PEG3* and *GRB10* DMRs, some patients were also displaying a combination of congenital abnormalities and/or developmental delay along with the phenotype associated with TNDM1 as noted in the previous study by the same lab. A further study showed two patients with TNDM1, both patients had hypomethylation at the *PLAGL1* DMR which was expected, but were also observed with hypomethylation at the *KvDMR* (Arima, 2005).

In BWS aberrant methylation is expected to be at the *KvDMR* and/or *H19* imprinted regions, though one study has associated MLID in up to 30% of BWS patients with hypomethylation at the *KvDMR*, although is less common in BWS patients with *H19* hypermethylation (Sanchez-Delgado et al., 2016a). NLRP family members NLRP2, NLRP5 and NLRP7 have also been associated with MLIDs in a term described as ‘maternal-effect’ mutations. As mentioned in the previous chapter, maternal NLRP2 germ-line mutation is associated with BWS in offspring with loss of methylation at the *KCNQ1OT1* and *MEST* DMRs (Meyer et al., 2009). NLRP5 is the most recent of the NLRP family to be involved with MLIDs: maternal NLRP5 mutations led to the loss of imprinting for a number of genes in various combinations, with some of the offspring displaying BWS- and SRS-like phenotypes (Docherty et al., 2015). NLRP7 is the most investigated of the NLRP family members involved with MLIDs: maternal NLRP7 mutations have been involved in BWS, affecting methylation at several different gDMRs and recurrent

hydatidiform mole, which shows normal biparental chromosomes but complete loss of maternal imprinted DMRs (Begemann et al., 2018; Mahadevan et al., 2014; Murdoch et al., 2006). These maternal effect genes have been investigated using mouse models, however, they have not provided insight into the mechanisms behind MLID as mutations at these genes result in early embryo lethality or cause limited imprinting alterations in the progeny (Mahadevan et al., 2017; Sparago et al., 2019; Tong et al., 2000).

3.1.5 Uniparental disomy

Many different models have been used in an attempt to understand genomic imprinting such as samples from imprinting disorder patients, UPD at a single chromosome or genome-wide UPD, *in vitro* mutations and/or treatment of cell lines, hydatidiform moles and uniparental embryonic stem cells (Allen et al., 1994; Barletta et al., 1997; Court et al., 2014; Devriendt, 2005; Ding et al., 2015; Elalaoui et al., 2014; Quenneville et al., 2011; Sharp et al., 2010). UPD is the inheritance of chromosome(s) from a single parent, it is observed in almost all types of imprinting disorders and was first described by Engel in 1980 (Eggermann et al., 2015b; Engel, 1980). UPD can be either isodisomy, in which the chromosome pair is a duplicate from one parent or heterodisomy in which the pair is heterozygous, but still from the same parent. UPD can arise due to 1) gamete complementation, 2) hetero-chromosomal substitution, 3) post-fertilization errors, 4) monosomy rescue causing chromosome duplication and 5) trisomic rescue followed by chromosomal loss (Engel, 1993; Spence et al., 1988). Genome-wide UPD is a very rare occurrence with under 20 cases being described in the literature since 2017 and only 3 of these cases being genome-wide maternal UPD. As expected, mUPD is observed with a maternal bias for methylation and expression of cognate imprinted genes. Interestingly these patients present with an SRS-like phenotype (Bens et al., 2017; Strain et al., 1995; Yamazawa et al., 2010). Genome-wide pUPD is against prediction as androgenetic embryos would be expected to incorrectly develop to an embryoblast and instead become a hydatidiform mole (Devriendt, 2005). Despite this, a handful of patients with genome-wide pUPD have been described in the literature, they observe a methylation and expression bias towards paternally imprinted genes and present with BWS-like phenotypes (Morales et al., 2009; Romanelli et al., 2011; Yamazawa et al., 2011).

3.1.6 The inefficiency of imprinting disorder testing

The majority of genomic imprinting syndromes rely on molecular tests to confirm the condition. While this may be appropriate for PWS, which is well-characterised with 3-4 abnormalities all focused within one region, it is just not enough for the remaining imprinting disorders. Many of the imprinting conditions especially SRS have a proportion of clinically diagnosed patients in which no underlying molecular defect could be found using standard tests such as MS-MLPA. This is disadvantageous to the patient as many imprinting syndromes have varying phenotypes depending on the molecular disturbance that is causing the disorder. Diagnosis based on clinical features could lead to a misdiagnosis of the patient and therefore preventing the patient from receiving the suitable treatment. With the increased use of genome-wide methylation tools, MLID is being seen as a common occurrence among many imprinting disorders and presents with varying phenotypes, combining the known clinical features involved within their respective imprinting disorders and additional phenotypes in some cases. Current diagnostic tests lack the power to accurately identify patients with imprinting disorders and the full extent of their molecular defects without multiple tests especially with the increased observation of MLIDs. The Illumina bead chip array has been described as a promising tool for molecular diagnostics (Grafodatskaya et al., 2016). Due to the increased occurrences of MLID and variable phenotypes liked with the disturbances it is imperative that a diagnostic test for imprinted disorders has the capacity to diagnose disturbances outside the disease-causing region. With this in mind, the genome-wide bead chip array would be a strong candidate. However, before the bead chip array can be utilised in such manner a set of normal ranges for each DMR will need to be correctly defined to be able to conclude what constitutes as a gain or loss of methylation at imprinted DMRs (Monk et al., 2018).

3.1.7 Characterisation of imprinted DMRs using Infinium arrays

A study by Hernandez *et al* (2018) previously characterized the imprinted regions using the 850k array and identified imprinted status at new genes using UPD patient data. The authors also went on to characterize normal and abnormal levels of methylation by comparing the genomic UPD patient 850k datasets to normal 850k leukocyte datasets similar to what was

carried out in this thesis. To carry this out Hernandez *et al* (2018) used ± 1SD from the control data as their statistical cut-off.

In the previous chapter, we observed that the methylation data at a subset of imprinted DMRs were non-normally distributed in human blood 450k datasets. With this mind, we believed it to be better to use the median along with the IQRs to create a statistical cut-off for normal ranges of methylation for each DMR. Further to this, we will also compare these cut-offs against numerous imprinted disorder patient datasets for PWS, AM, SRS and BWS to detect abnormal methylation in patient datasets rather than just UPD patients as seen in the Hernandez study. Due to the lack of 850k datasets for imprinting disorder patients, we will carry out this study using 450k array datasets.

The overall aim of this chapter is to characterise normal ranges of methylation in the 450K array using the median and IQRs. The normal ranges will then be tested for the ability to detect molecular defects at imprinted regions in order to provide a more powerful diagnostic tool for describing multiple molecular defects and the presence of MLID in one test.

3.2 Aims and objectives

Aims

Can we use the 450k array to:

1. Identify normal methylation at imprinted regions
2. Categorise a set of testable DMRs
3. Score epimutations at the selected DMRs in specific syndromes

Objectives:

1. Score methylation at 2-3 well-characterised imprinted regions in the blood
2. Test normal range for these imprints in several tissues to assess variability among somatic tissue
3. Apply this to all known imprints to see how many can be scored like this
4. Define a criterion for normal variability at DMRs in 100 blood control samples
5. With the use of a defined criterion is it possible to score abnormal methylation defects to diagnose Angelman syndrome, Prader–Willi syndrome, Silver Russell syndrome and Beckwith Wiedemann syndrome patient 450k samples?

3.3 Methods

R-studio was used to generate bed files of 450k array data sets to be uploaded and analysed using USE-GALAXY as described in chapter 2. Normal somatic tissue 450 datasets used in this chapter was blood (GSE55491), saliva (in house dataset), muscle, liver and brain (GSE52578). Genome-wide UPD patient samples for maternal and paternal methylation controls (GSE52578). 100 healthy blood samples were used to create a normal reference range (GSE73103). Finally, imprinting disorder datasets used in this chapter was Angelman syndrome (GSE78956), Prader-Willi syndrome (GSE78956), Beckwith-Wiedemann syndrome (GSE78956 and GSE95488) and finally Silver-Russell Syndrome (GSE55491). These datasets while all very detailed and furthered our understanding of imprinting disorders, they use extremely complex bioinformatic approaches which would not be applicable to a diagnostic setting without highly experienced training. Further from this, only Bens et al (2016) describe what they believed to be abnormal changes of methylation but this may not be a suitable method in order to effectively detect epimutations in patients.

3.4 Results

3.4.1 Establishing testable DMRs

To test the 450k array a detailed collection of imprinted DMRs was assembled using the literature, as gDMRs are more stably kept as lifelong marks they make up the majority of the list along with some sDMRs that have been relatively well described (Table 1). This list as generated using the genome location as defined by Court et al (2011) and Woodfine et al (2011). While both studies thoroughly investigated imprinting DMRs, Woodfine et al (2011) carried out investigations by validating 50 pyrosequencing assays for imprinting DMRs across numerous somatic tissues. However, a shortcoming of this paper is that it did not confirm the methylation bias on the alleles with other methods. Court et al (2014) furthered the investigations into genomic imprinting DMRs with the use of genome-wide methylation infimum bead chip array on a number of somatic tissues and genome-wide UPD patients. Using this approach Court *et al* was able to show several novel regions with parental specific methylation, though they did not investigate normal variation at the imprinted DMRs and had a small number of somatic tissues as controls. Using these two papers a list of coordinates was generated to investigate normal variability of imprinting DMRs across a large pool of samples. Then using this variability to create a methylation index (MI) to identify epimutations to diagnose potential imprinting disorders. The initial list was produced based on the establishment of the DMR and a selection of novel DMRs which showed promise. Some DMRs in the original formulation of the list showed methylation bias but only contained 1-5 probes lowering its statistical power, however, are described to be disturbed in some imprinting disorders so all DMRs with low probe coverage were kept as to not discriminate, though this low coverage will be considered during the discussion.

Table 1

Name	Chr	Start	End	MO	Est	Probe (n) 450k	Probe (n) 850k	Source	ID	Reference
PPIEL	chr1	40024626	40025540	M	UNK	4	4	Court	PHP1B	(Rochtus et al., 2016)
DIRAS3_ex2	chr1	68512505	68513486	M	S	8	8	Both	PHP1B	(Begemann et al., 2018; Fontana et al., 2018; Sanchez-Delgado et al., 2016a)
DIRAS3	chr1	68515433	68517545	M	GL	21	23	Both	PHP1B	
ZDBF2	chr2	207114583	207136544	P	UNK	8	14	Both	IUGR/MLID	(Monteagudo-Sánchez et al., 2019)
NAP1L5	chr4	89618184	89619237	M	GL	15	13	Both	MLID	(Bak et al., 2016; Sanchez-Delgado et al., 2016a)
FAM50B	chr6	3849082	3850359	M	UNK	25	24	Court	MLID	(Bens et al., 2016)
PLAGL1	chr6	144328078	144329888	M	GL	17	15	Court	TNDM1	(Mackay et al., 2005)
IGF2R	chr6	160426265	160427561	M	UNK	4	5	Both	N/A	N/A
GRB10g	chr7	50848726	50851312	M	GL	9	10	Both	SRS	(McCann et al., 2001)
PEG10	chr7	94284759	94287960	M	GL	72	68	Both	SRS	(Hannula-Jouppi et al., 2013)
MEST	chr7	130130122	130134388	M	GL	62	54	Both	SRS	(Hannula-Jouppi et al., 2013)
HTR5A	chr7	154862719	154863382	M	UNK	6	8	Court	SRS	(Hannula-Jouppi et al., 2013)
CXORF56_ ERLIN2	chr8	37604992	37606088	M	UNK	7	7	Court	N/A	N/A
TRAPPC9	chr8	141108147	141111081	M	UNK	8	9	Court	N/A	N/A
INPP5F	chr10	121577530	121578727	M	GL	6	7	Both	MLID	(Maeda et al., 2014)
H19	chr11	2018812	2024740	P	GL	51	43	Both	BWS/SRS	(Bartholdi et al., 2009; Catchpoole et al., 1997)
IGF2_DMR2	chr11	2153834	2155112	P	S	10	10	Woodfine	MLID	(Maeda et al., 2014)
IGF2_DMR0	chr11	2168333	2170145	P	S	1	2	Woodfine	MLID	
KvDMR1	chr11	2719948	2722259	M	GL	33	26	Both	BWS	(Smilinich et al., 1999)
RB1	chr13	48892341	48895763	M	GL	13	13	Both	SRS	(Prickett et al., 2015)
MEG3	chr14	101290524	101293978	P	S	33	35	Both	TS	(Temple et al., 1991)
MEG8	chr14	101370741	101371419	M	S	1	1	Court	TS	(Temple et al., 1991)
MAGEL2	chr15	23892425	23894029	M	UNK	6	9	Court	PWS /AM	(Clayton-Smith and Laan, 2003; Zahova and Isles, 2018)
NDN	chr15	23931451	23932759	M	S	8	10	Court	PWS/AM	
SNRPN_ intraCpG29	chr15	24671872	24672679	M	UNK	4	3	Court	N/A	N/A
SNRPN_ intraCpG40	chr15	25017924	25018886	M	UNK	4	4	Court	N/A	N/A
SNRPN_1	chr15	25068564	25069481	M	UNK	9	8	Court	N/A	N/A

SNRPN_2	chr15	25093008	25093829	M	UNK	4	4	Court	N/A	N/A
SNRPN_3	chr15	25123027	25123905	M	UNK	6	5	Court	N/A	N/A
SNURF	chr15	25200004	25201976	M	GL	7	10	Both	PWS/AM	(Clayton-Smith and Laan, 2003; Zahova and Isles, 2018)
IGF1R	chr15	99408496	99409650	M	UNK	7	7	Court	MLID	(Begemann et al., 2018)
ZNF597	chr16	3481801	3482388	M	GL	2	3	Court	Upd16mat	(Inoue et al., 2018)
ZNF597_ NAA60	chr16	3492828	3494463	P	S	12	15	Court	Upd16mat	
ZNF331	chr19	54057086	54058425	M	S	16	30	Court	IUGR	(Petre et al., 2018)
PEG3	chr19	57348493	57353271	M	GL	37	36	Both	MLID	(Baple et al., 2011; Boonen et al., 2013)
MCTS2P_ HM13	chr20	30134663	30135933	M	GL	10	8	Both	N/A	
BLCAP_NNAT	chr20	36148604	36150528	M	S	37	37	Both	MLID	(Sanchez-Delgado et al., 2016a)
L3MBTL	chr20	42142365	42144040	M	GL	26	26	Both	MLID	(Bak et al., 2016; Sanchez-Delgado et al., 2016a)
GNAS	chr20	57414039	57418612	P	UNK	23	26	Court	MLID	(Bak et al., 2016; Baple et al., 2011; Bliek et al., 2009; Rochtus et al., 2016)
NESP-AS_GNAS-AS1	chr20	57425649	57428033	M	S	66	59	Court	MLID	
GNAS_XL	chr20	57428905	57431463	M	GL	6	7	Both	MLID	
GNAS_1A	chr20	57463453	57467939	M	S	62	58	Both	MLID	
NHP2L1	chr22	42077774	42078873	M	UNK	8	8	Court	MLID	(Bak et al., 2016; Rochtus et al., 2016)

Table 3.1: Imprinted DMRs used for investigation. The table shows the location within the genome, parent of origin, the establishment of the mark, probe coverage based on the 450k and 450k array, source of the DMR location and any associated disorders linked to loss of imprinting at the respective DMR and the reference is a source linking the DMR to the respective disorder**.** M, Maternal; P, Paternal; Unknown, Methylation and/or Marks establishment unknown; PHP1B, pseudohypoparathyroidism type 1B; MLID, Multilocus imprinting disturbance; IUGR, intrauterine growth restriction; TNDM1, Diabetes Mellitus, Transient Neonatal 1; SRS, Silver Russell Syndrome; PWS, Prader-Willi syndrome; AM Angelman syndrome; BWS, Beckwith-Wiedemann syndrome; Upd16mat, maternal uniparental disomy of chromosome 16; TS, Temple Syndrome; GL. Germ-line; S, Somatic; UNK, Unknown; ID, Imprinting Disorders; MO, Methylation Origin; Chr, Chromosome; Est, Established.

3.4.2 Categorising imprinted methylation with 450k

To initially look to see whether the methylation bead chip arrays can differentiate between imprinted DMRs/ICRs and other methylated controlled regions, 3 well-described DMRs (*H19*, *KvDMR* and *SNURF*) were compared to hypermethylated and hypomethylated CGIs at the promoter regions of *DAZL* and *MTHFR* respectively in the blood (Figure 3.3). It was clear that the methylation at the imprinted DMRs investigated was indicative of imprinted DMRs (35-65%) while the *DAZL* and *MTHFR* promoter CGI's showed clear hypermethylation and hypomethylation, respectively, in both the 450k (A) and EPIC (B) array. The EPIC was analysed due investigate whether greater probe coverage at the regions resulted in any changes of variability at the DMR (Hernandez Mora et al., 2018). As well as greater probe coverage the EPIC array also covers the IG-DMR not covered by the 450k array as shown in the next chapter (Figure 4.5). Due to lack of Imprinted disorder datasets for the 850k array, the 450k array will be used for further analysis.

Figure 3.3

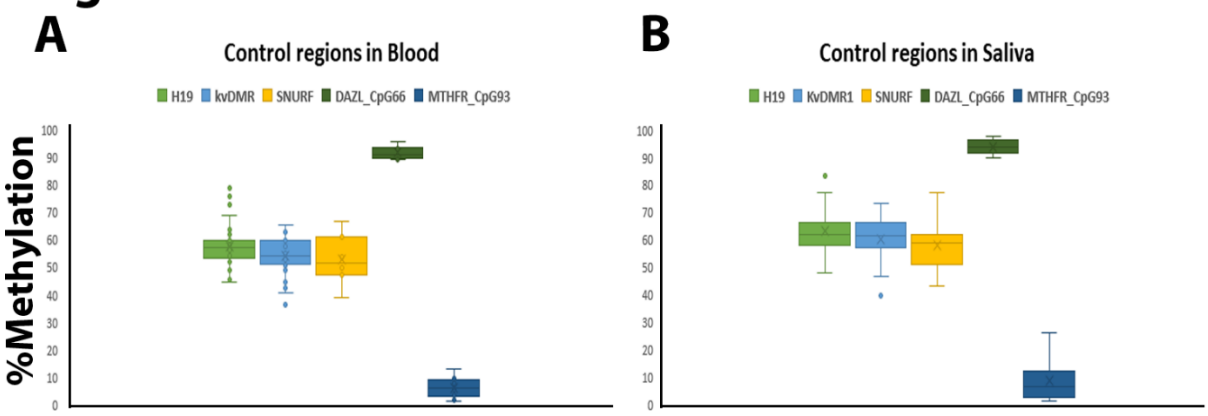


Figure 3.3: Ability of the array to detect well-established imprinted DMRs in blood and saliva tissue. Average methylation of probes found within the genomic coordinates for the *H19* (51 probes (450k) 43 probes (EPIC)), *KvDMR* (33 probes (450k) 26 probes (EPIC)) and *SNURF* (7 probes (450k) 10 probes (EPIC)) DMRs with *DAZL* (Chr3:16646340-16647237) (8 probes (450k) 12 probes (EPIC)) and *MTHFR* (Ch1:11865387-11866780) (15 probes (450k) 19 probes (EPIC)) promoter region CGI's as positive and negative controls of common methylation control regions in (A) 5 healthy peripheral blood 450k samples and (B) 23 healthy Saliva 850k samples. X; Mean, Line; Median, Box; interquartile range (IQR), Whiskers; IQR*1.5, Circles; probe outliers. Using the Mann Whitney U test (MWU) all imprinted DMRs tested were significantly different ($p<0.001$) to both *DAZL* and *MTHFR* promoter CGIs.

Figure 3.4

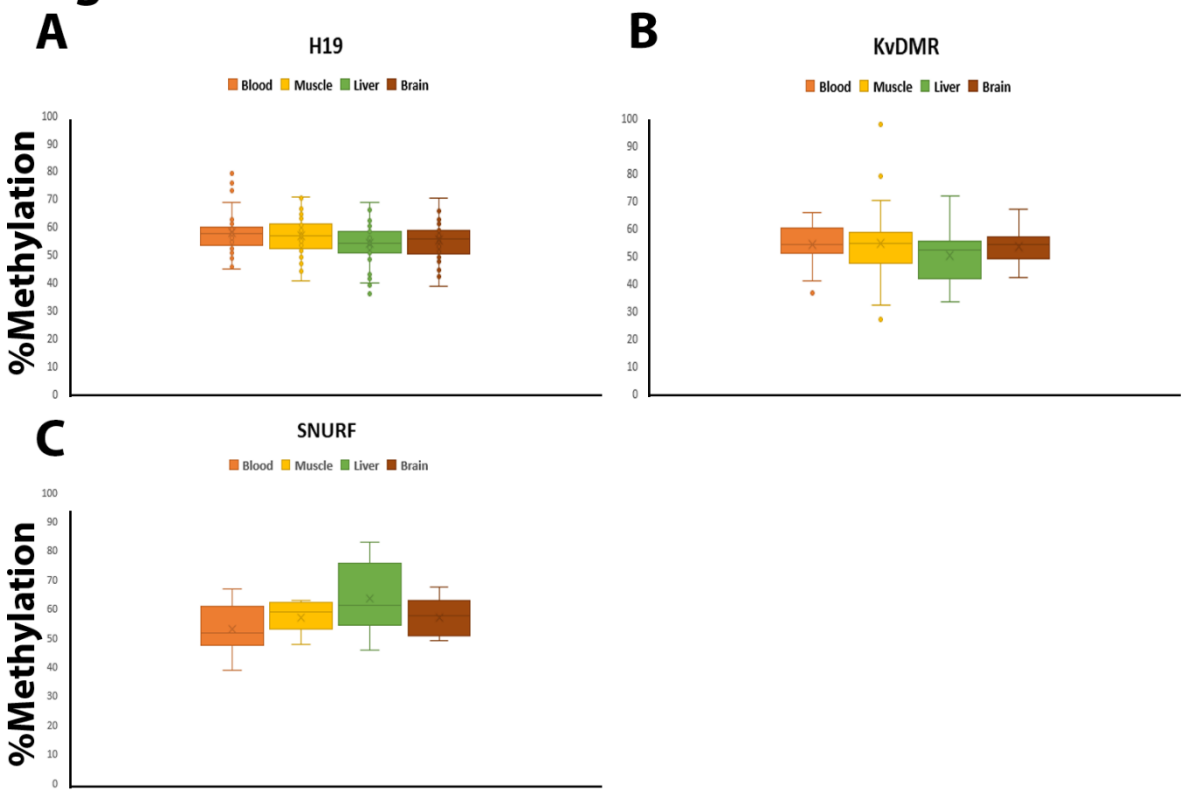


Figure 3.4: Variability at well-established DMRs across different somatic tissues Average methylation of probes found within the genomic coordinates for the (A) *H19* (51 probes), (B) *KvDMR* (33 probes) and (C) *SNURF* (7 probes) DMRs in 450k blood (n=5), muscle (n=1), liver (n=1) and brain (n=1). The box represents the interquartile range (IQR), the whiskers are IQR*1.5; the X marks the mean and a line represents the median; circles are outlier probes. MWU significance in comparison to blood demonstrated by *p < 0.05; **p < 0.01; ***p < 0.001.

To further understand the ability of the 450k array to categorise imprinted DMRs the variability across different somatic tissues was examined. As expected, the average methylation fell within the expected imprinted DMR range of 35-65% (Figure 3.4). As expected, normal methylation was observed across the imprinted DMRs tested, with only the liver sample appearing as statistically different when compared to blood for the imprinted DMRs *H19* and *KvDMR*. Although sample numbers were low for other somatic tissues, these results nevertheless suggested that in blood and saliva at least (two readily accessible tissues) the imprinted DMRs kept within tightly defined bounds.

We now wished to test if the 450k could be used to differentiate abnormal from normal methylation patterns at these DMRs. To do so, 450k datasets from the blood of rare cases of genome-wide maternal and paternal UPD were utilised (Court et al., 2014). The GSE52578 dataset was from the detailed beadchip analysis from Court et al (2014), the authors investigated a large collection of imprinted DMRs in different somatic tissues and genome-wide UPD patients allowing for novel DMRs to be discovered. However, for categorising imprinted DMRs the authors lacked sample numbers so would not show the true value of methylation-based on sample variability. When compared to the control blood dataset, the genome-wide UPD patients' samples showed the expected methylation bias i.e. the maternally methylated *H19* DMR was hypermethylated and hypomethylated in the mUPD and paternal UPD (pUPD) respectively and vice versa in the paternally methylated DMRs *KvDMR* and *SNURF* (Figure 3.5A-C). As a control, the CGIs at the promoter region of *DAZL* and *MTHFR* remain hyper- and hypomethylated respectively, though small differences (<10%) can be detected (Figure 3.5 D-E). Therefore, proof of concept has been shown, that the 450k can detect methylation levels indicative of normal methylation levels in control blood with the ability to differentiate between normal and abnormal samples due to unique methylation changes at the imprinted DMRs.

Figure 3.5

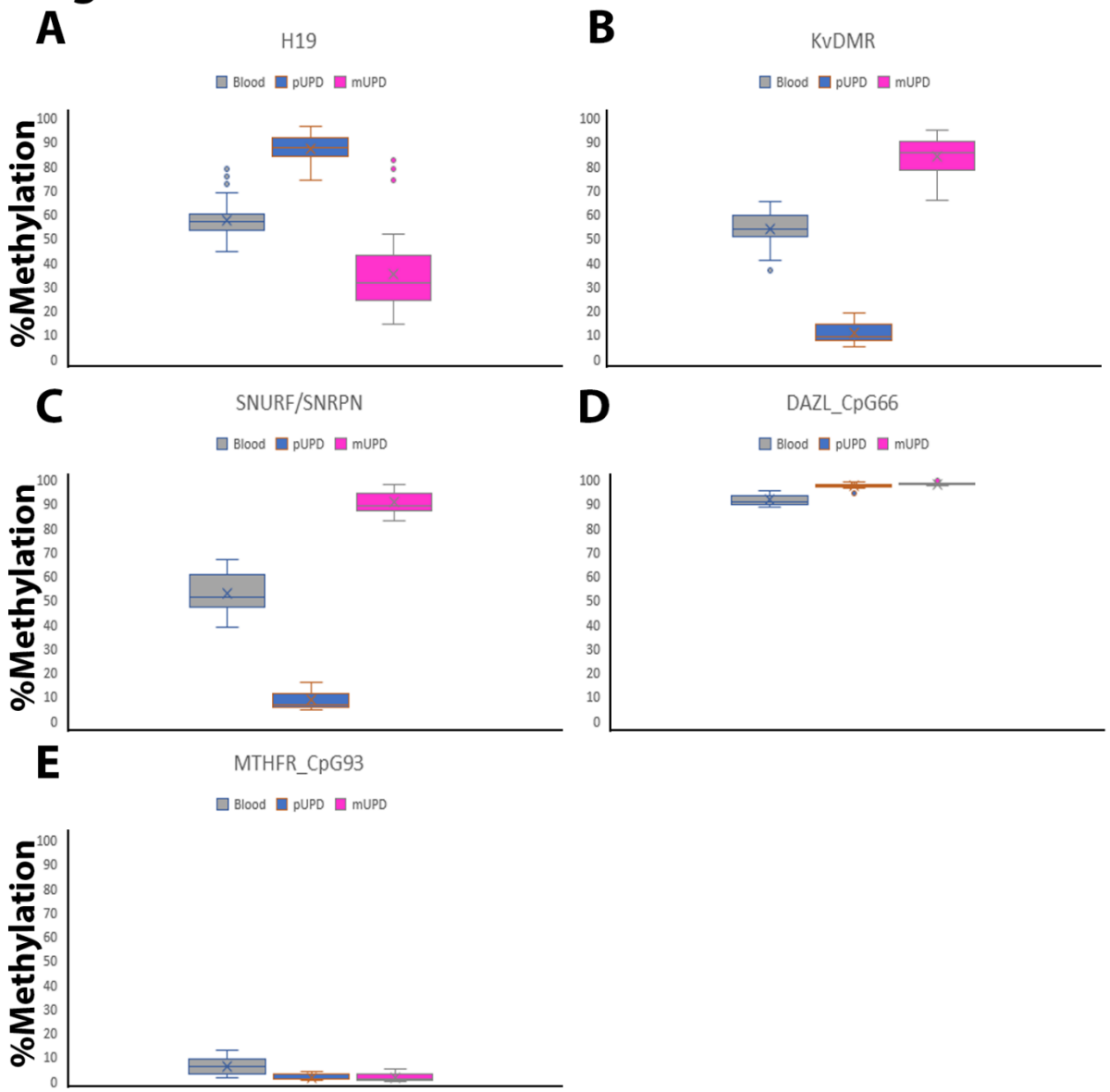


Figure 3.5: The 450K array can differentiate between normal and abnormal methylation in tissues with known uniparental bias. Utilising 450k array data for genome-wide UPD, it is possible to see distinct changes of methylation at the Imprinted DMRs near the (A) *H19,* (B) *KvDMR* and (C) *SNURF* loci in comparison to a normal blood control. To ensure that the changes of methylation between the genome-wide UPD samples and blood were unique to the Imprinted DMRs methylation was examined at the heavily methylated CGI at the (D) *DAZL* promoter region and the lightly methylated CGI at the (E) *MTHFR* promoter region. pUPD, paternal uniparental disomy; mUPD, maternal uniparental disomy. Box and whiskers as before, MWU *p < 0.05; **p < 0.01; ***p < 0.001.

3.4.3 Application of concept to a larger set of imprinted DMRs

Since it was established that aberrant methylation could be detected at the *H19, KvDMR* and *SNURF/SNRPN* DMRs in the control UPD samples, the same method was applied to all of the imprinted DMRs described in Table 3.1. Many of these DMRs were not as well described as the previously tested DMRs. Therefore, to determine the normal ranges for each maternally methylated DMR the datasets GSE73103 (Blood) and GSE52578 (Muscle, Liver and Brain) were utilised to score methylation at imprinted DMRs across different somatic tissues, though only blood was used for further downstream analysis in order to match tissue from the imprinted disorder datasets (Figure 3.6A). Similarly, this was carried out for the paternally methylated DMRs (Figure 3.6B).

Figure 3.6

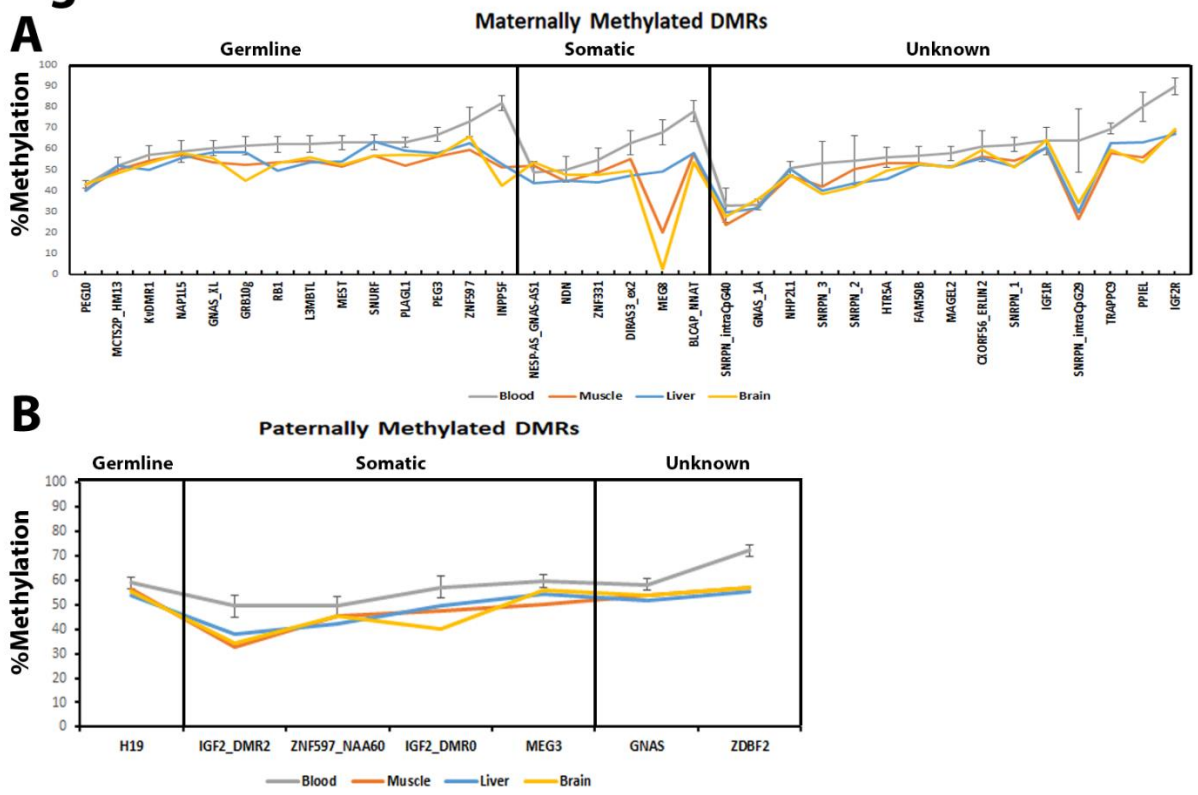


Figure 3.6: Characterising methylation at Imprinted DMRs across somatic tissues. (A)Variability of methylation at paternally methylated Imprinted DMRs in different somatic tissues. (B) Variability of methylation at maternally methylated Imprinted DMRs in somatic tissues. The average methylation was calculated using all probes within the DMR for each individual blood control sample. The average DMR methylation for each sample was then averaged and the median methylation at each DMR was used to plot the line graph with error bars representing IQR. Only 1 sample was used for Muscle, Liver and Brain so mean methylation for each DMR was plotted.

Imprinted DMRs were excluded if methylation levels were above 70% or less than 40% a similar exclusion criteria carried out by (Hernandez Mora et al., 2018). Further to this if the interquartile range (IQR) was over 10% these DMRs were also excluded from any further testing due to having methylation not indicative of a DMR or an IQR suggestive of a highly variable region as indicated by the red boxes in Figure 3.7 A and B. The excluded regions are listed in Fig. 3.7C: a number of these are novel regions defined by a single study (Court et al., 2014) or further studies from within the same group and so did not have extensive support for being imprinted. While the methylation bias is also detectable here, the authors only used a small sample size of somatic tissues, within a large collection of normal samples these regions appear to be highly variable which would make it difficult to detect abnormal methylation in patient samples.

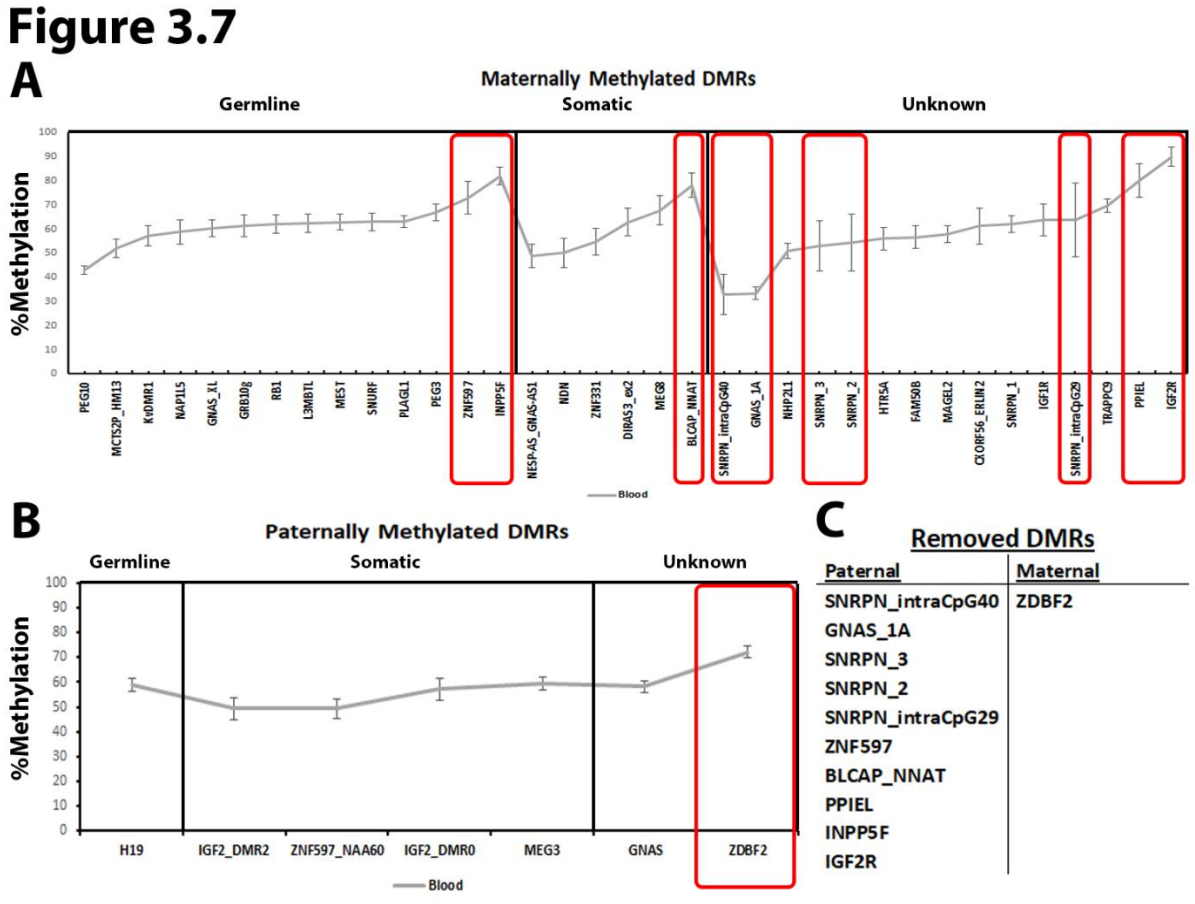


Figure 3.7: Characterising methylation at Imprinted DMRs in control blood. (A)Variability of methylation at (A) maternally-methylated and (B) paternally-methylated imprinted DMRs. The average methylation was calculated using all probes within the DMR for each individual blood control sample and the median methylation from 100 samples plotted, with error bars representing IQR. Red boxes indicate DMRs excluded from further testing, listed in (C).

As a further criterion for using the array to score methylation, we compared methylation in the normal blood samples to that of the genome-wide UPD patient samples: if the median methylation values in either of the UPD samples were within the IQR of the control blood samples, as indicated by the red boxes in Figure 3.8A, that DMR was also excluded from further testing. Using these criteria excludes any DMRs which may complicate further results due to the inability to determine normal and abnormal methylation results. The DMRs remaining show clear methylation bias in both pUPD and mUPD samples within a methylation range of 40-70% (Figure 3.8A and B).

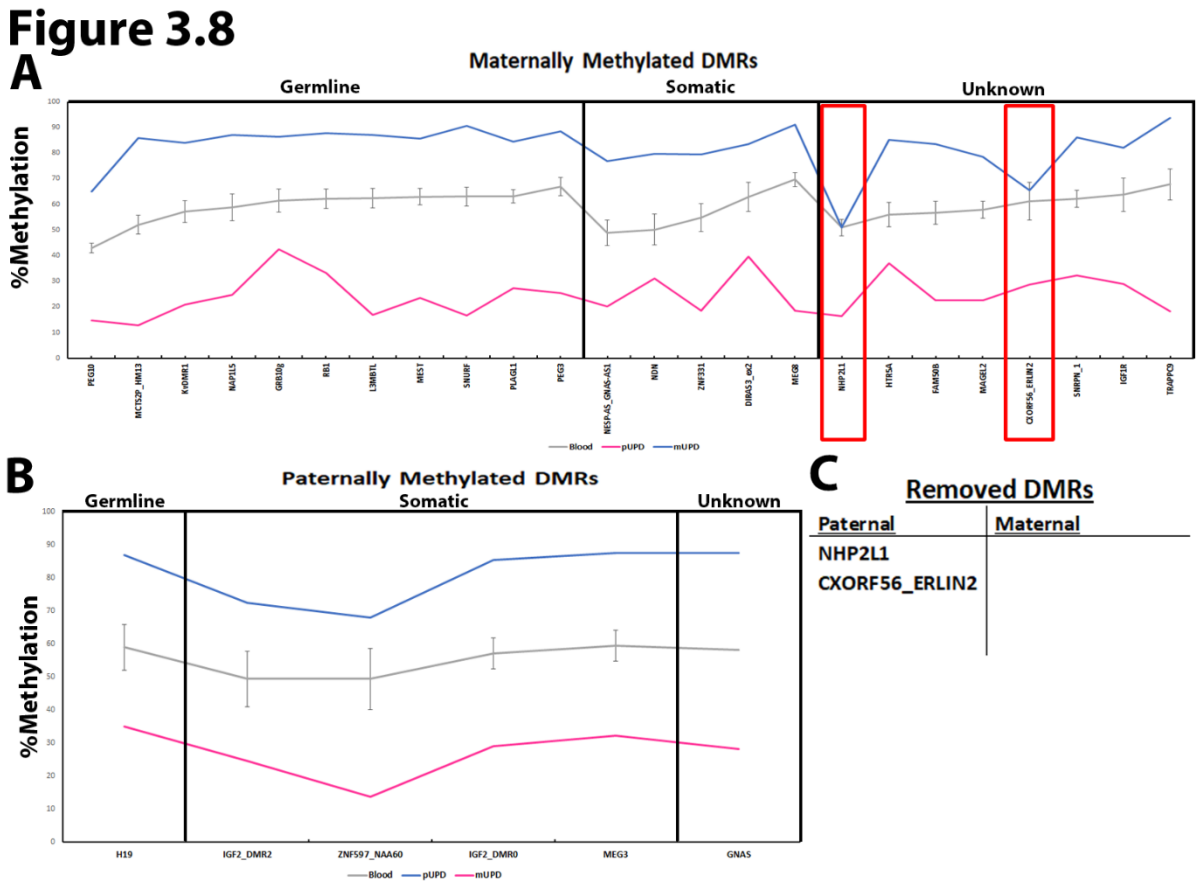


Figure 3.8: Investigating abnormal methylation at imprinted DMRs. (A) Methylation at paternally methylated Imprinted DMRs in control blood samples and paternal and maternal genomewide uniparental disomy patients. **(B)** Methylation at maternally methylated Imprinted DMRs in control blood samples and paternal and maternal genomewide uniparental disomy patients. The average methylation was calculated using all probes within the DMR for each individual blood control sample. The average DMR methylation for each sample was then averaged and the median methylation at each DMR was used to plot the line graph with error bars representing IQR. Red boxes indicate DMRs excluded from further testing and are listed in **(C).**

Once this clear set of testable Imprinted DMRs had been established, they were used to investigate methylation abnormalities in imprinting disorder patients. To facilitate this, a methylation index (MI) was formulated based on median methylation of each individual DMR across 100 normal patient blood samples and using the Interquartile rule to find outliers was used as a criterion to define gains and losses of methylation outside the normal range so that each MI was uniquely based on the intervariable methylation of each distinctive imprinted DMR associated with single-locus disorders. For multi-locus imprinting disorders (MLID) an adjusted MI was used with an additional +/-5% to increase stringency and therefore the likelihood that the change of methylation is involved in MLID as many of these regions are not functionally described or previously observed in the literature to be involved with MLID.

3.4.4 Ability of the array to accurately identify Prader-Willi and Angelman syndrome patients

As the molecular diagnosis of PWS is well described (see chapter Introduction), this was the first to be investigated using the array to determine if the MI criteria would be sufficient to accurately score patients with epigenetic changes. To do so, we availed of a published dataset from GSE78956 which was generated during a study carried out by Bens et al (2016) investigating the phenotypical spectrum and methylation defects within MLID. This was a good collection of imprinting disorder patients containing 37 patients' samples of varying imprinting disorders and compared them against 39 control patient samples to test if the 450k array could detect methylation disturbances in patient samples outside the normal disease-causing locations. The authors here described hypomethylation and hypermethylation values with changes over 10% from the normal blood samples. This was a highly valuable collection due to the variety and number of imprinting disorders examined using the infimum bead chip array especially for the work planned to be carried out in this chapter. RStudio along with the RNbeads package was utilised to download this dataset from the GEO expression omnibus as described in chapter 2. Table 2 shows the result of processing the dataset from Bens *et al* (2016), concentrating on Chromosome 15p etc around the defined imprinted genes SNURF/SNRPN. Three PWS and three AM patients were available and their methylation values was compared to a defined MI by using the median methylation seen across the same regions in 100 control blood samples. PWS patients showed maternal specific methylation at the

imprinting cluster on chromosome 15 while the AM patients appeared with paternal specific methylation. The abnormal methylation discovered here had a 100% match to the previously diagnosed epimutations. The MI described above, identified all tested DMRs within this imprinted cluster as abnormal across all patients with hypomethylation or hypermethylation in PWS and AM respectively. As this dataset also contained sample data for TS and TNDM, these imprinting disorders were also tested using the MI. However, the dataset only contained one patient sample for both TS and TNDM so the data was not shown here but can be seen in Appendix Table A1 and A2.

Table 2

Imprinted DMR	Est	MO	Chr	Start	End	Blood	PWS 1	PWS 2	PWS 3	AM 1	AM 2	AM 3	MI Upper	MI Lower
IGF1R	U	M	chr15	99408496	99409650	63.94	51.15	56.16	56.13	51.67	57.79	50.35	81.88	45.73
MAGEL2	G	M	chr15	23892425	23894029	58.39	81.33 *	82.99 *	82.86 *	23.92 *	20.38 *	15.48 *	64.86	51.14
NDN	G	M	chr15	23931451	23932759	50.99	73.73*	71.6 *	71.07 *	28.53 *	28.45 *	24.43 *	62.15	37.88
SNRPN_1	G	M	chr15	25068564	25069481	62.48	80.53*	83.94 *	83.79 *	32.89 *	29.22 *	22.69 *	69.2	55.78
SNURF	G	M	chr15	25200004	25201976	63.73	93.94*	92.87 *	92.84 *	15.07 *	11.42 ***	4.09 *	70.84	56.4

Table 3.2: Median methylation across imprinted DMRs in Angelman and Prader-Willi syndrome patient samples. Methylation values with red underlined text showed gains of methylation relative to the methylation index (MI) for that DMR, while blue underlined text shows loss of methylation relative to the MI. PWS patients show clearly the expected hypermethylation at the chromosome 15 imprinted cluster while AM shows clear hypomethylation at the same region. The blood sample displays the mean methylation across 100 control blood samples, while disease samples represent individual cases. MWU in comparison to blood *p < 0.05; **p < 0.01; ***p < 0.001. MO, Methylation Origin; Chr, Chromosome; Est, Established; G, Germ-line; M, Maternal, P; Paternal; MI, Methylation Index; DMR, Differentially methylated region; AM, Angelman; PWS, Prader-Willi syndrome.

3.4.5 Using the 450k array for diagnostic testing in Beckwith-Wiedemann syndrome

BWS has a slightly more complex molecular diagnosis than that of PWS and AM but contains clear molecular aberrations involved with the disorder. To test this, two datasets from the GEO expression omnibus GSE95488 and GSE78956 from two studies carried out by Bens et al (2016) as was used for the PWS and AM patients and a case study of 3 BWS patients carried out by Brzezinski et al (2017). Brzezinski et al (2017) evaluated the tumour surveillance programme within children with *KvDMR* epimutations, this dataset extended the number of BWS samples we could investigate including a mosaic patUPD patient. Like the dataset used for AM and PWS patients, the GSE95488 dataset was downloaded and processed using R studio and the RNbeads package. The expected molecular defects at chromosome 11 were observed in all BWS patient according to the MI, 1 patient had methylation indicative of Pat11UPD due to hypermethylation at the *H19* DMR and hypomethylation at the *KvDMR* (BWS1). Two patients (22%) were observed with hypermethylation at *H19* (BWS 4 and 5) while the remaining 6 were identified with hypomethylation at *KvDMR* (66%) (Table 3). The methylation abnormalities identified by the MI were seen to be significant through testing by MWU. Further to this, the MI had a 100% match to what was previously described for the patient samples.

Table 3

Imprinted DMR	Est	MO	Chr	Start	End	Blood	BWS 1	BWS 2	BWS 3	BWS 4	BWS 5	BWS 6	BWS 7	BWS 8	BWS 9	MI Upper	MI Lower
MEG8	G	M	chr14	101370741	101371419	68.89	59.3	63	63.6	62.11	55.33	59.44	58.7	57.05	55.37	85.09	51.033
H19	G	P	chr11	2018812	2024740	59.2	64.74*	59.03	57.3	86.81*	86.73*	55.82	60.15	54.97	59.73	63.95	54.29
KvDMR1	G	M	chr11	2719948	2722259	57.98	45.12*	26.11*	28.22*	55.64	55.7	24.56*	21.72*	16.73*	43.4*	66.04	49.04

Table 3.3: Table methylation across imprinted DMRs in Beckwith-Wiedeman syndrome patient samples. Methylation across imprinted DMRs associated with the respective diseases, methylation values with red underlining text showing gains of methylation outside the defined methylation index (MI) while blue underlined text shows loss of methylation outside the defined MI. All patients show the expected epimutations either displaying methylation values suggesting Pat11UPD, due to hyper and hypomethylation at the *H19* and *KvDMR* respectively, or singular epimutations at either *H19* or *KvDMR*. The blood sample displays the median methylation of the average methylation of the DMR across 100 control blood samples, while disease samples display mean methylation at the DMR as individual samples. Mann Whitney u significance in comparison to blood demonstrated by *p < 0.05; **p < 0.01; ***p < 0.001. MO, Methylation Origin; Chr, Chromosome; Est, Established; G, Germ-line; M, Maternal, P; Paternal; MI, Methylation Index; DMR, Differentially methylated region; BWS, Beckwith-Wiedeman syndrome.

3.4.6 Potential for diagnostic testing in Silver-Russell syndrome with the 450K array

SRS is one of the more complicated imprinting disorders, so it provides an opportunity to fully test the ability of the 450k array and the defined MI to detect aberrant methylation. The GSE55491 data set generated by Prickett et al (2015) carried out a genome-wide methylation analysis on a large collection of SRS patients with detailed patient data. They showed that some epimutations were detectable in the 450k datasets in patients that had no molecular disturbances identified in previous diagnostic tests. The focus of this paper was to test a novel methodology to analyse DMR changes genome-wide describing some novel DMR locations across their samples. As above the dataset was downloaded and processed through the use of Rstudio and the RNbeads package. The MI defined here was able to confirm all SRS patients that were diagnosed by Pricket *et* al (2015) including those that were not able to be previously diagnosed using current standard diagnostic tests. Despite this, 50% of the cases (n=18) showed no methylation disturbances at the expected DMRs such as mat7UPD or *H19*. Of these, abnormal methylation suggestive of mat7UPD was observed in 3 of the SRS patients (SRS Patient 3, 8 and 15) while another 6 were observed with hypomethylation at the *H19* DMR (SRS Patient 4, 10, 12, 13, 14 and 17) (Table 4). Again, each epimutation identified through the use of the MI were seen to be statistically significant through the use of MWU. Further to this, each patient-matched the epimutation that was described by Pricket *el al* (2015), however, the methodology used to 'score' the DMRs defined in this thesis is less bioinformatically complex and would be easier to be carried out by a non-specialised bioinformatician.

Table 4

Imprinted DMR	Est	MO	Chr	Start	End	Blood	SRS 1	SRS 2	SRS 3	SRS 4	SRS 5	SRS 6	SRS 7	SRS 8	SRS 9	MI Upper	MI Lower
NAP1L5	G	M	chr4	89618184	89619237	60.09	59.55	60.53	59.83	59.84	60.86	59.85	65.68	59.61	59.65	69.21	48.67
GRB10g	G	M	chr7	50848726	50851312	61.75	63.17	60.17	87.69*	59.26	59.21	63.2	63.36	87.53*	61.54	70.59	52.79
HTR5A	G	M	chr7	154862719	154863382	56.08	52.6	50.8	71.05*	54.67	58.27	55.03	57.92	71.43*	51.65	65.35	46.47
MEST	G	M	chr7	130130122	130134388	63.44	59.9	57.39	91.26*	59.4	60.47	59.85	60.75	92.81*	58.66	69.33	56.4
PEG10	G	M	chr7	94284759	94287960	43.14	42.85	40.78	70.86*	41.73	43.09	43.34	43.3	70.28*	42.27	46.7	39.41
H19	G	P	chr11	2018812	2024740	59.2	59.81	59.41	58.25	49.88*	60.87	59.47	56.66	59.77	59.74	63.95	54.29
Imprinted DMR	Est	M-O	Chromosome	Start	End	Blood	SRS 10	SRS 11	SRS 12	SRS 13	SRS 14	SRS 15	SRS 16	SRS 17	SRS 18	MI Upper	MI Lower
NAP1L5	G	M	chr4	89618184	89619237	60.09	57.57	59.08	58.95	58.79	59.73	54.11	59.86	64.87	55.15	69.21	48.67
GRB10g	G	M	chr7	50848726	50851312	61.75	58.07	58.97	59.82	61.27	62.27	84.03*	62	61.5	57.46	70.59	52.79
HTR5A	G	M	chr7	154862719	154863382	56.08	54.48	54.07	52.8	58.32	54.42	72.45*	54.15	55.58	58.28	65.35	46.47
MEST	G	M	chr7	130130122	130134388	63.44	56.18*	58.51	55.97*	58.05	60.52	92.98*	60.25	60.28	53.32*	69.33	56.4
PEG10	G	M	chr7	94284759	94287960	43.14	39.57	41.31	41.49	40.52	41.49	70.62*	41.94	41.8	38.15	46.7	39.41
H19	G	P	chr11	2018812	2024740	59.2	47.09*	59.21	41.02*	43.76*	41*	61.21	59.01	36.09*	54.58	63.95	54.29

Table 3.4: Table methylation across imprinted DMRs in Silver-Russell syndrome patient samples. Methylation across imprinted DMRs associated with the respective diseases, methylation values with red underlining text showing gains of methylation outside the defined methylation index while blue underlined text shows loss of methylation outside the defined methylation index. The 50% of patients present with the expected methylation abnormalities such as mat7UPD or *H19* hypomethylation. The blood sample displays the median methylation of the average methylation of the DMR across 100 control blood samples, while disease samples display mean methylation at the DMR as individual samples. Mann Whitney u significance in comparison to blood demonstrated by *p < 0.05; **p < 0.01; ***p < 0.001. MO, Methylation Origin; Chr, Chromosome; Est, Established; G, Germ-line; M, Maternal, P; Paternal; MI, Methylation Index; DMR, Differentially methylated region; SRS, Silver-Russell syndrome.

3.4.7 Multilocus imprinting disturbances

A current restriction to the current diagnostic testing is the inability to diagnose MLID that may be involved with the patients. Few studies have described the need for diagnostic tests to diagnose MLID in imprinting disorder patients, with the beadchip array being described as a potential candidate to carry this out, however, one issue is what loci's should be considered as MLID and to what extent do changes of methylation result in MLID (Eggermann et al., 2014; Eggermann et al., 2016; Sanchez-Delgado et al., 2016a). In this chapter, I aim to further investigate these issues by defining a more stringent MI than that was used for identifying disease-causing loci. This increased stringency should help removed potential false positives increasing the likelihood that the identified target is involved in MLID and could warrant further investigations. This was carried out as sDMRs have variable methylation levels on different alleles, the increased range will also help rule out this variability, helping confirm true loss or gain of methylation from the hypermethylated/hypomethylated alleles, respectively. As sDMRs normally only affect the expression of the nearby imprinted gene rather than the entire cluster aberrant methylation may result in phenotypical changes to the disease respective to the gene/sDMR affected.

MLID in AM and PWS is not commonly seen as it is in other imprinting disorders (Sanchez-Delgado et al., 2016a). It is clear from the results for this the AM and PWS patients from GSE78956, that the disturbance is primarily localised to the large imprinting region on chromosome 15 and is not suggestive of chromosomal 15 UPD, as these patients do not appear to have abnormal methylation at the *IGF1R* DMR. However, 66% of AM and PWS patients were observed with possible MLID targets outside the expected regions. 33% of AM and PWS patients appeared with hypomethylation at the *PLAGL1* DMR just outside the defined MI (Table 5). *PEG3* DMR hypomethylation was observed in 33% of PWS patients and all AM patients, like the disturbances at the *PLAGL1* DMR the *PEG3* DMR hypomethylation was just outside the defined MI. Finally, the *RB1* DMR was observed to be hypomethylated across all patients with multiple methylation disturbances, though similar to the previous DMRs they only just fell outside the defined MI.

Out of the imprinting disorders examined here, BWS is most commonly found to present with MLID (Azzi et al., 2009). Using the datasets GSE95488 and GSE78956 possible MLID targets were indicated in 88% (8/9) of the BWS patients with BWS7 being the only exception. The somatic *IGF2_DMR0* was seen to be also hypermethylated in 100% of BWS patients with *H19* hypermethylation as the primary epimutation. 33% of patients were observed with extreme MLID with 3+ DMRs affected outside the disease-causing primary epimutation of *KvDMR*. Most of the disturbances outside the disease-causing regions were observed to be hypomethylated and just below the defined MI. However, extreme LOM (~10-20%) was observed in BWS3 at both *DIRAS3* DMRs and the *IGF1R* and *L3MBTL1* DMRs (Table 6).

GSE55491 a dataset generated by Prickett *et* al (2015) was used to investigate potential targets of MLID and explain the SRS phenotype in patients with no primary epimutation defined. This was done with the use of the expanded DMR list. MLID was only observed in SRS patients with *H19* hypomethylation. 66% (3/6) of SRS patients with the primary *H19* DMR epimutation were observed with a further methylation disturbance at the *IGF2_DMR0* DMR. SRS 10 was observed with hypomethylation at the *PLAGL1* DMR just below the defined lower MI. SRS12 was observed with multiple disturbed DMRs with a LOM below the defined MI for *MEG8* (2.63%) and *IGF1R* (27.76%) DMRs. SRS9 a patient with no primary epimutation showed an unexpected LOM at the *MEG3* somatic DMR and hypermethylation at the *MEG8* gametic DMR suggesting maternal methylation bias at this region or possibly the maternal chromosome 14 UPD (mUPD14). mUPD14 is a disease-causing aberration observed in TS and interestingly the methylation pattern observed at chromosome 14 in SRS9 is the same pattern observed in the TS patient sample in Appendix Table A1 (Temple et al., 1991). A further SRS patient (SRS 18) with no primary epimutations was observed with a 2.77% LOM outside the defined MI at the *SNURF* DMR (Table 7).

Table 5

Imprinted DMR	Est	MO	Chr	Start	End	Blood	PWS 1	PWS 2	PWS 3	AM 1	AM 2	AM 3	MI Upper	MI Lower
PLAGL1	G	M	chr6	1.44E+08	1.44E+08	63.1	52.81	59.89	55.64	53.09	55.64	54.65	73.48	53.34
RB1	G	M	chr13	48892341	48895763	62.17	49.04	53.36	49.65	44.95	50.71	48.1	75	49.65
MAGEL2	G	M	chr15	23892425	23894029	57.9	81.33	82.99	82.86	23.92	20.38	15.48	64.86	51.14
NDN	G	M	chr15	23931451	23932759	50.16	73.73	71.6	71.07	28.53	28.45	24.43	62.15	37.88
SNRPN_1	G	M	chr15	25068564	25069481	62.11	80.53	83.94	83.79	32.89	29.22	22.69	69.2	55.78
SNURF	G	M	chr15	25200004	25201976	63.05	93.94	92.87	92.84	15.07	11.42	4.09	75.84	56.4
PEG3	G	M	chr19	57348493	57353271	66.92	54.37	59.32	57.05	54.66	56.8	54.23	78.95	55.01

Table 3.5: Table investigating the incidence of MLID in Angelman and Prader-Willi syndrome patient samples. Legend as before (Table2), Imprinted DMRs with abnormal methylation falling outside the primary disease-causing regions in AM and PWS patients as highlighted by outside borders. 2/3 of AM and PWS patients are observed with MLID outside the primary disturbed location at chromosome 15. Regions flagged by the MI are consistent at the same regions across both disorders which are *PLAGL1*, *RB1* and *PEG3.*

Table 6

Imprinted DMR	Est	MO	Chromosome	Start	End	Blood	BWS 1	BWS 2	BWS 3	BWS 4	BWS 5	BWS 6	BWS 7	BWS 8	BWS 9	MI Upper	MI Lower
DIRAS3	G	M	chr1	68515433	68517545	60.22	46.92	55.87	31.49	54.64	45.28	48.04	54.66	51.69	47.26	71.58	47.78
DIRAS3_ex2	G	M	chr1	68512505	68513486	62.48	49.25	47.7	28.64	52.13	46.59	47.95	52.26	50.17	44.5	79.03	46.03
FAM50B	U	M	chr6	3849082	3850359	58.02	51.2	35.64	41.62	53.48	51.97	51.1	55.35	50.97	47.75	71.41	43.06
PLAGL1	G	M	chr6	144328078	144329888	63.66	53.26	59.02	55.89	60.29	56.41	57.4	62.38	53.92	53.08	73.48	53.34
MEST	G	M	chr7	130130122	130134388	63.44	54.95	57.47	43.96	59.18	56.37	54.59	59.95	55.23	52.52	74.33	51.4
H19	G	P	chr11	2018812	2024740	59.2	64.74	59.03	57.3	86.81	86.73	55.82	60.15	54.97	59.73	63.95	54.29
IGF2_DMR0	S	P	chr11	2168333	2170145	57.18	62.2	63.5	56.6	71.93	81.24	51.08	55.56	51.06	57.89	71.15	43.39
KvDMR1	G	M	chr11	2719948	2722259	57.98	45.12	26.11	28.22	55.64	55.7	24.56	21.72	16.73	43.4	66.04	49.04
RB1	G	M	chr13	48892341	48895763	62.48	53.22	55.8	51.82	52.58	53.53	52.76	55.39	47.87	47.02	75	49.65
IGF1R	U	M	chr15	99408496	99409650	63.94	55.46	54.91	35.94	59.16	57.23	81.52	55.99	52.14	45.79	81.88	45.73
SNURF	G	M	chr15	25200004	25201976	63.73	56.01	58.29	55.23	55.34	55.02	52.66	56.84	50.99	47.18	75.84	51.4
PEG3	G	M	chr19	57348493	57353271	67.4	56.67	60.46	58.7	59.18	52.39	54.81	60.29	54.29	53.14	78.95	55.01
ZNF331	G	M	chr19	54057086	54058425	54.8	45.03	48.21	44.66	43.22	41.04	45.85	45.72	40.78	37.69	70.81	39
L3MBTL	G	M	chr20	42142365	42144040	63.37	54.86	58.45	32.78	55.88	59.47	54.44	54.66	54.61	50.61	75.1	49.74
NESP-AS_GNAS-AS1	S	M	chr20	57425649	57428033	48.81	48.02	30.31	40.26	52.08	42.54	46.18	47.21	50.39	42.01	64.02	34.17

Table 3.6: Table investigating the association of MLIDs in Beckwith-Wiedeman syndrome patient samples. Legend as before (Table3), Imprinted DMRs with abnormal methylation falling outside the primary disease-causing regions in BWS patients as highlighted by outside borders. All patients with the exception of BWS patient 7 were flagged with MLID outside the primary epimutation. This was found throughout a number of different DMRs, but a clear pattern of hypermethylation was observed between the gametic and somatic DMRs *H19* and *IGF2_DMR0*. *DIRAS3* and *PEG3* DMRs were observed to be disturbed outside the primary epimutation in ½ of the BWS patients. BWS patients with the primary epimutation at the *KvDMR* appears to be more prone to numerous MLIDs outside the primary epimutation.

Table 7

Imprinted DMR	Est	MO	Chr	Start	End	Blood	SRS 1	SRS 2	SRS 3	SRS 4	SRS 5	SRS 6	SRS 7	SRS 8	SRS 9	MI Upper	MI Lower
PLAGL1	G	M	chr6	1.44E+08	1.44E+08	63.66	59.88	58.78	58.24	54.22	59.01	57.52	57.8	57.92	57.16	73.48	53.34
GRB10g	G	M	chr7	50848726	50851312	61.75	63.17	60.17	87.69	59.26	59.21	63.2	63.36	87.53	61.54	75.59	47.79
HTR5A	G	M	chr7	1.55E+08	1.55E+08	56.08	52.6	50.8	71.05	54.67	58.27	55.03	57.92	71.43	51.65	70.35	41.47
MEST	G	M	chr7	1.3E+08	1.3E+08	63.44	59.9	57.39	91.26	59.4	60.47	59.85	60.75	92.81	58.66	74.33	51.4
PEG10	G	M	chr7	94284759	94287960	43.14	42.85	40.78	70.86	41.73	43.09	43.34	43.3	70.28	42.27	51.7	34.41
H19	G	P	chr11	2018812	2024740	59.2	59.81	59.41	58.25	49.88	60.87	59.47	56.66	59.77	59.74	68.95	49.29
IGF2_DMR0	S	P	chr11	2168333	2170145	57.18	59.1	55.7	60.3	51.2	60.6	63.5	46.5	61.8	58.1	71.15	43.39
MEG3	S	P	chr14	1.01E+08	1.01E+08	59.58	59.21	56.3	56.46	57.93	60.17	59.64	59.79	59.32	23.74	69.79	48.9
MEG8	G	M	chr14	1.01E+08	1.01E+08	68.89	58.5	61.2	61.1	58	60.3	61.7	60.2	58.2	98.6	85.09	51.03
IGF1R	U	M	chr15	99408496	99409650	63.94	55.86	52.99	56.83	51.61	56.07	57.01	57.21	58.73	53.57	81.88	45.73
SNURF	G	M	chr15	25200004	25201976	63.73	55.11	54.26	54.37	55.51	56.46	55.96	53.76	56.09	51.96	75.84	51.4

Imprinted DMR	Est	MO	Chr	Start	End	Blood	SRS 10	SRS 11	SRS 12	SRS 13	SRS 14	SRS 15	SRS 16	SRS 17	SRS 18	MI Upper	MI Lower
PLAGL1	G	M	chr6	1.44E+08	1.44E+08	63.66	52.25	53.58	56.52	56.34	56.98	55.16	56.91	57.99	54.44	73.48	53.34
GRB10g	G	M	chr7	50848726	50851312	61.75	58.07	58.97	59.82	61.27	62.27	84.03	62	61.5	57.46	75.59	47.79
HTR5A	G	M	chr7	1.55E+08	1.55E+08	56.08	54.48	54.07	52.8	58.32	54.42	72.45	54.15	55.58	58.28	70.35	41.47
MEST	G	M	chr7	1.3E+08	1.3E+08	63.44	56.18	58.51	55.97	58.05	60.52	92.98	60.25	60.28	53.32	74.33	51.4
PEG10	G	M	chr7	94284759	94287960	43.14	39.57	41.31	41.49	40.52	41.49	70.62	41.94	41.8	38.15	51.7	34.41
H19	G	P	chr11	2018812	2024740	59.2	47.09	59.21	41.02	43.76	41	61.21	59.01	36.09	54.58	68.95	49.29
IGF2_DMR0	S	P	chr11	2168333	2170145	57.18	43.8	58.2	38.9	40.8	44.1	59.1	61.5	42.2	53.6	71.15	43.39
MEG3	S	P	chr14	1.01E+08	1.01E+08	59.58	55.61	57.41	58.45	58.06	58.53	57.28	58.34	59.03	53.17	69.79	48.9
MEG8	G	M	chr14	1.01E+08	1.01E+08	68.89	61.3	63.6	48.4	59.2	56.9	54.2	60.4	55.8	56.2	85.09	51.03
IGF1R	U	M	chr15	99408496	99409650	63.94	49.86	55.86	17.97	52.47	56.79	52.13	58.59	60.09	57.86	81.88	45.73
SNURF	G	M	chr15	25200004	25201976	63.73	52.94	55.11	55.59	52.63	55.97	53.56	55.39	53.56	48.63	75.84	51.4

Table 3.7: Table methylation across imprinted DMRs in Silver-Russell syndrome patient samples. Legend as before (Table4), Imprinted DMRs with abnormal methylation falling outside the primary disease-causing regions in SRS patients as highlighted by outside borders. Similar to the observation in BWS patients, almost all SRS patients that contain primary epimutations at the *H19* DMR was also observed with hypomethylation at the somatic *IGF2_DMR0*. Disturbances outside the primary epimutation was limited to mostly 1 other DMR with the exception of SRS 12 which was flagged with MLID at the *MEG8* and *IGF1R* DMRs. Patients with methylation patterns suggestive of mat7UPD appeared with no MLIDs. SRS patient 9 did not contain any epimutations associated with SRS but disturbances were detected at the *MEG3, MEG8* and SNURF DMRs using the MI.

3.5 Discussion

How genomic imprinting and its related disorders are investigated is currently lacking consistency and can often be confusing. To shed light on these issues papers have suggested 1) a nomenclature system when discussing Imprinted DMRs, 2) the need for multilocus testing in imprinting disorders, 3) a clear definition for gains and losses of methylation at imprinted DMRs and 4) development of a consistent and reproducible molecular method for testing (Monk et al., 2018; Poole et al., 2013). The extant techniques used in molecular diagnosis are presently not sufficient for imprinting disorders, given the complexity behind imprinting disorders, especially SRS which contains multiple regions with molecular abnormalities being linked to the disease phenotype and the increasing number of investigations showing MLID and its association with severity of phenotype (Poole *et al.*, 2013). At present, molecular diagnosis is carried out by investigating the primary epimutations involved with the imprinting disorder and/or chromosomal abnormalities such as UPD or deletions. However, many molecular abnormalities are undiagnosed in imprinting disorders and locus-specific assays are not able to detect MLID. In some cases current methods have led to misdiagnosis due to the overlap in clinical features, which may cause further complications due to incorrect treatment not targeting the actual disorder (Fontana et al., 2018; Luk, 2016; Petit et al., 2012; Wakeling et al., 2016). Due to these reasons, it is important to identify the molecular defects as well as any other disturbances that may be altering the phenotype efficiently and effectively in the syndrome in order to fully understand the underlying disorder and treat accordingly. In this chapter, the 450k array was examined to address some of the issues with the current diagnostic methods.

The 450k bead chip array is one of the most cost-effective and least bioinformatically complex tools for whole-genome DNA methylation profiling (Dedeurwaerder et al., 2011). The 450k array has proven to be an invaluable approach for investigating genomic imprinting and its related disorders by being used as a tool for identifying novel DMRs at/near imprinted genes, as well as identifying further molecular disturbances associated with imprinting disorders and profiling patients with MLID (Court et al., 2014; Prickett et al., 2015; Rezwan et al., 2015). In this chapter, I utilised publicly available datasets and processed them with R and then analysed them using a bespoke GALAXY workflow CandiMeth (Thursby et al., 2020) developed

by the Walsh lab which does not require high-grade bioinformatic skills. The DMRs were thoroughly tested to get a collection which shows normal methylation with little variability between the 100 control blood samples yet showed clear differences between control and disease state in blood samples. This was an important step to simplify the test, as many of the regions excluded are defined as DMRs in the literature (Court et al., 2014; Sharp et al., 2010; Woodfine et al., 2011). In this investigation, even though methylation bias is observed in the genome-wide UPD samples for these DMR, the regions were excluded as diagnostic sites as they were observed to differ from canonical imprinted DMR by having excess levels of methylation and/or large variability between control blood samples.

The array with the defined MI was able to identify the expected methylation abnormalities which had been previously diagnosed/described in the source articles. Not only was it possible to identify DMRs with epimutations but it was able to indicate UPD in the respective patients due to parental-specific methylation at multiple DMRs. The results from the AM and PWS patients agrees with the literature in that the chromosome 15 imprinting region was associated with the disorders, though a larger dataset would be required to fully investigate the use of the defined MI in these imprinting disorders (Buiting et al., 2016; Butler et al., 2016). The epimutations identified in BWS patients were also in line with the literature in which hypomethylation at the *KvDMR* was the leading cause of BWS, followed by hypermethylation at the *H19* DMR, then chromosome 11 UPD (Ibrahim et al., 2014; Luk, 2017; Weksberg et al., 2010). Finally, the results for the SRS patients was in accord with the literature with hypomethylation at the *H19* DMR being the most common epimutation of SRS, followed by mat7UPD and a large portion of SRS patients with no identifiable underlying molecular cause (Azzi et al., 2015; Fokstuen and Kotzot, 2014; Õunap, 2016).

As no common DMRs have been associated with MLID for any imprinting disorder and the expanded list of DMRs are not characterised as well as the disease-causing DMRs, an extra +/-5% was applied to the MI, to increase confidence that this was a disturbance and not just variability. Further to this, many DMRs on the expanded list have little information on their functionality such as the extent to which changes of methylation are required to change the expression of genes controlled by the DMR, this is an important factor to consider when investigating MLID purely from methylation status as changes of methylation outside the MI

may be detected though not necessarily lead to a physiological effect. MLID in AS and PWS does not appear too common with only two AM cases being described with MLID and no PWS patients described with MLID (Sanchez-Delgado et al., 2016a). The method applied in this study observed LOM outside the expected imprinting region of chromosome 15, at 3 other DMRs on/near the imprinted genes *PLAGL1, RB1* and *PEG3* in both AM and PWS. Both *RB1* and *PEG3* were previously described with LOM in AM patients, further suggesting that these two DMRs may play a role in MLID in these disorders (Baple et al., 2011). BWS has been observed to present with multiple imprinting disturbances outside the expected regions, more than any other imprinting disorder examined in this study (Azzi et al., 2009). Here we show a large proportion of BWS patients (88%) presented with many hypomethylated regions outside the expected DMRs, this was extremely high in comparison to the literature (Maeda et al., 2014). Despite this, all of the regions identified as possible targets for MLIDs have been previously described in the literature (Bliek et al., 2009; Docherty et al., 2014; Sanchez-Delgado et al., 2016a). The BWS patients which were observed with hypermethylation at the *H19* DMR were also observed with concurrent hypermethylation at the nearby somatic DMR *IGF2_DMR0*, though this was not observed in the UPD patient. On the other hand, SRS patients with *H19* DMR hypomethylation was observed with nearby hypomethylation at the somatic *IGF2_DMR0.* Hypomethylation at the *PLAGL1* and *IGF1R* DMRs have been previously shown for MLID in SRS, while the *MEG8* DMR to our knowledge has not been previously described as associated with SRS (Sanchez-Delgado et al., 2016a).

Apart from identifying the epimutation causing the disorder and possible MLID targets, through the use of the array and defined MI, we may have detected a misdiagnosed patient. SRS9 was diagnosed without molecular abnormalities and presents with clinodactyly (curvature of a digit) and growth restriction with a negative SDS birth weight and height and weight at assessment (Prickett et al., 2015). Considering the patient information and the maternal specific methylation pattern observed at the *MEG3* and *MEG8* DMRs this would suggest Temple syndrome. Temple syndrome presents with many clinical features overlapping with SRS, including but not limited to pre-and postnatal growth alterations and clinodactyly. A recent study by Luk (2016) has described a case in which a one-year-old boy was initially misdiagnosed with SRS when actually suffering from Temple syndrome, when

MEG3 DMR hypomethylation was detected (Habib et al., 2019; Kagami et al., 2017). SRS 18 only showed hypomethylation at the *SNURF* DMR, though not to the extent seen in AS.

While the 450K and the defined MIs showed clear benefits for use as a diagnostic tool for imprinting there are some minor restrictions to this method. A number of the imprinted DMRs have little probe coverage in the 450k array, while some have none at all such as the IG-DMR on chromosome 14. The 100 blood samples allowed for assessment of variability in mean methylation at the DMRs, but the low probe coverage at some DMRs may mean this is not truly representative of methylation across the region. This could be improved by using the 850k (EPIC) array which increases the probe coverage from 716 to 789 across the imprinted DMRs as well containing two probes at the IG-DMR not originally covered by the 450k array (Hernandez Mora et al., 2018). While the MI for the MLID was larger than that of the known aetiological DMRs, targets identified may not have functional effects and would require further investigations to implicate these regions in MLID. A number of improvements could help reduce the current limitations of this approach, such as the use of the 850k array to investigate genome-wide methylation and using a larger sample size of imprinting disorder patients some without a previous molecular diagnosis. The ability to confirm transcriptional changes for the imprinting disorder patients through PCR or RNA-seq techniques would also allow for the DMRs outside the disease-causing DMRs to be better characterised and further our understanding of MLIDs. It would also be beneficial to apply this method to more novel imprinting disorders such as TS and TNDM etc in order to prove that the concept applies to all imprinting disorder (Sanchez-Delgado et al., 2016a). Further use of the beadchip array that was not explored here is copy number variation (CNV) analysis, which can be done by using probe signal ratios and is built into a number of the newer or updated methylation beadchip array analysis pipelines such as WateRmelon and RnBeads 2.0 (Assenov et al., 2014; Pidsley et al., 2013). Use of the array to estimate CNV at the imprinted regions could help to explain the defects in SRS patients with no canonical methylation disturbances (Feng et al., 2013).

3.6 Summary

The 450k has proven itself to be an invaluable tool for methylation studies of the epigenetic phenomena imprinting. In this chapter, the array was investigated as a tool for molecular diagnosis in selected imprinting disorders. The 450k array along with the methylation index (MI) described here was seen to be an accurate tool providing a more detailed genome-wide assessment than current molecular diagnostic methods. The 450k could diagnose aberrant methylation in previously tested patients and differentiate between UPD and epimutations associated with imprinting disorders. Further to this it also presented with an added benefit of being able to identify possible targets for MLID outside the expected region for all of imprinting disorder patients as well as describing a patient which was likely misdiagnosed. The results here show that the beadchip array (450k or 850k) could be utilised as a diagnostic tool alongside scoring of clinical features to diagnose patients with imprinting disorders, this would prove to be especially beneficial in cases in which SRS is suspected due to the complexity of this disorder. A better categorisation and understanding of the disease provided by the beadchip array would be beneficial for accurate treatment and reduction of misdiagnosis for imprinting disorder patients. Not only would the increased usage of the beadchip array have a benefit to the patient if the datasets were also available for research purposes it could be used to further investigate the mechanism behind imprinting and imprinting-like disorders as well as the methylation disturbances behind MLIDs. As the understanding of the methylation mechanism behind the less described DMRs is expanded the MI can be altered to determine the sensitivity of the DMR to losses or gains of methylation.

Chapter 4- DMR cross-talk in DNMT1 depleted hTERT-1604 human fibroblastic cell line

4.1 Introduction

4.1.1 Establishment of imprinting DMRs and mechanism of action

Imprinting DMRs can be established in both the germline and after fertilisation in somatic tissues which are known as germline DMRs (gDMR) or somatic DMRs (sDMR) respectively as described previously (Nechin et al., 2019). gDMRs are defined by their establishment during gametogenesis and maintenance as a lifelong mark, this establishment and maintenance is reliant on the action of DNMTs. gDMRs are mostly methylated on the maternal allele are usually found within a cluster of imprinted regions controlling their expression (Cheong et al., 2015). A further characteristic of gDMRs are the open chromatin and close chromatin histone marks found on the active and inactivate alleles respectively. Further to this maternally methylated DMR's are usually found intragenic generally corresponding to promoters often of lncRNAs, while paternally methylated gDMRs are intergenic and function as insulators or enhancers (Edwards and Ferguson-Smith, 2007). sDMRs are defined by their establishment outside the germ line and may be tissue-specific and depend on the presence of a gDMR (Ferguson-Smith, 2011). sDMRs are relatively rare and correlate with gene promoters and transcription factor binding sites, higher-order chromatin structure, chromatin modifications and lncRNAs involved with silencing nearby imprinted genes in cis (Barlow and Bartolomei, 2014; Monk et al., 2019).

As discussed in the introduction to this thesis the H19 and IGF2 DMRs are a well-described model of the hierarchical function of the gDMR. A further imprinting cluster on chromosome 14 appears to display a hierarchy interaction in which loss of the gDMR *IG-DMR* by deletions is associated with hypermethylation of the sDMR *MEG3* (Beygo et al., 2015; Kagami et al., 2010). Furthering this a study observed that transcription from the *MEG3* promoter was essential for establishing methylation at another sDMR at *MEG8* (Beygo et al., 2017). A similar interaction was also observed at the imprinting cluster on chromosome 20 in which the gDMR found at the *GNAS* loci would result in hypermethylation at the *NESP* sDMR suggesting that the gDMR is required for its establishment (Elli et al., 2018). However, conflicting this the

imprinting status of somatic DMR at the *Igf2r* locus in mouse appears to be independent of the gDMR (Stöger et al., 1993).

Considering the conflict we wished to investigate if in a system where methylation levels had been lowered, this affected both gDMRs and sDMRs equally, or whether effects were more marked on one and if so, was there a concomitant effect on the other. Preliminary examination of methylation in a human cell line model derived by the lab suggested that the latter was the case, so I set out to investigate this in more detail. For this, I chose three gDMR/sDMR pairs based on those which had the most normal levels of methylation in the model cell line. I will first describe these three loci, then outline the approaches to be taken. The three imprinted genes that I will investigate in this study are Growth Factor Receptor Bound Protein 10 (*GRB10*), Small nuclear ribonucleoprotein polypeptide N (*SNRPN*) and nucleotide-binding, leucine-rich repeat-containing proteins Family Pyrin Domain Containing 2 (*NLRP2*). The gDMRs at *GRB10* and *SNRPN* have been well characterised in both mouse and human showing functional importance (Blagitko et al., 2000; Nicholls and Knepper, 2001). We hypothesise that changing methylation at the gDMR will affect methylation at the sDMR. The alternative is that methylation at one will not affect methylation at the other.

4.1.2 Methylation controlled expression of SNRPN

SNRPN is a paternally expressed imprinted gene that is a part of the small nuclear ribonucleoprotein (SmN) family, *SNRPN* is located at the PWS locus on chromosome 15 (Figure 4.1). SNRPN is expressed from a bicistronic transcript which also encodes for SNRPN Upstream Reading Frame (SNURF) (Gray et al., 1999). The function of the genes at this imprinted cluster is poorly understood, however, the SmN family has been found in the mouse brain and heart where it is involved with pre-mRNA processing (Bervini and Herzog, 2013). Leff *et al* (1992) reported the maternal imprinting of the mouse *Snrpn* gene as well as its homology to the human gene and concluded that *SNRPN* could be a candidate gene for Prader-Willi syndrome (Leff et al., 1992). It was later reported that the expression of the human *SNRPN* gene was imprinted using Angelman and Prader-Willi patients (Glenn et al., 1993). A further study by Glenn *et al* (1996) reported preferential methylation at the transcription start site of *SNRPN* on the repressed maternal allele which was later described

in mouse as DMR1, this showed high resemblance between human and mouse (Glenn et al., 1996; Shemer et al., 1997). The whole cluster at this region is largely imprinted with a number of genes being paternally expressed due to imprinting and several maternally methylated DMRs close by and within the *SNRPN* locus itself, however, the DMR located at *SNURF-SNRPN* promoter is the gDMR and functionally defined as the ICR (Figure 4.1) with Bielinska *et al.*, (2000) showing that loss of this region in both human and mouse leads to maternal imprinting on both alleles (Court et al., 2014; Horsthemke and Buiting, 2006; Sharp et al., 2010). This ICR also known as the PWS-IC is established during oogenesis and drives the epigenetic imprinted control of the remaining DMRs (Chamberlain and Lalande, 2010b).*ATP10a* and *UBE3A* are the only maternally expressed genes in this region (Herzing et al., 2001; Rougeulle et al., 1997). The ATP10a gene contains a somatically acquired DMR which has been observed to be present in the placental tissue (Woodfine et al., 2011). Loss of function of the *UBE3A* gene results in AS which is the paternal imprinted disorder related to this region. However, it is unclear which gene is associated with PWS which is the maternal imprinted disorder for this region, it has been shown that the paternally expressed SNORD116 can result in a PWS-like phenotype (Buiting, 2010).

Figure 4.1

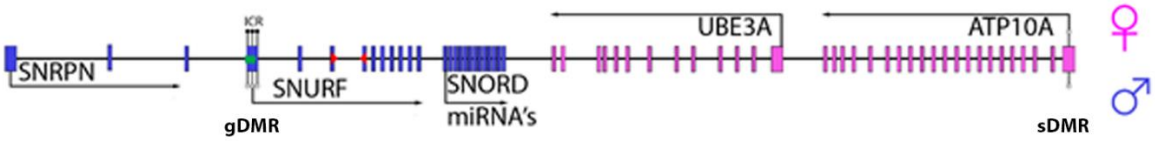


Figure 4.1: Methylation and expression in the human imprinted genes *SNRPN*. Most imprinted genes are found in clusters and are regulated by an Imprinted control region (ICR) or differentially methylated region (DMR). (A) The *SNURF/SNRPN* DMR Controls the paternal expression of *SNURF*, *SNRPN* and *SNORD* miRNAs and the maternal expression of *ATP10A* and *UBE3A*. The *ATP10a* DMR is hypomethylated in most tissues, though in placenta methylation levels consistent with that of a DMR can be seen. Key unchanged from figure 1.5.

4.1.3 Methylation-controlled expression of GRB10

GRB10 codes for an adapter protein on chromosome 7 that belongs to the Grb7 family of adaptor molecules, GRB10 interacts with receptor proteins and has been associated with metabolism, cell proliferation and apoptosis (Holt and Siddle, 2005). Mouse GRB10 knockout models have exhibited overgrowth in many organs and tissues as well as an increase of social dominance suggesting that GRB10 is important for growth regulation as well as affecting adult

behaviour (Charalambous et al., 2003; Garfield et al., 2011). *GRB10* does not have an imprinted disorder linked with aberrant methylation at the DMR, however, patients suffering from Silver-Russell Syndrome (SRS) have been associated with maternal UPD 7 (mUPD7), but the data is merely suggestive that the maternal diploidy of chromosome 7 causing overexpression of maternal *GRB10* may be the cause of the SRS symptoms, due to protein playing a role in growth regulation (Prickett et al., 2015). Blagitko *et al* (2000) showed that GRB10 expression was highly tissue- and isoform-specific as seen in Figure 4.2. In human GRB10 is biallelically expressed in most adult tissue from the major promoter. However, the imprinted expression is highly isoformic and tissue-specific with paternal expression being observed in the fetal brain from the brain-specific promoter which contains the maternally methylated gDMR, while maternal expression is only observed in skeletal muscle and placental trophoblast from the major promoter which contains the sDMR which is unmethylated on both alleles in these tissues. The imprinted maternal expression in skeletal muscle and placental trophoblast is yet to be determined with chromatin and methylation marks being ruled out by previous investigations (Blagitko et al., 2000; Monk et al., 2009). The mouse Grb10 is expressed in a maternal-specific manner in most tissues except the brain where paternal expression is observed. Interestingly the methylation pattern is consistent with that seen at the human locus, the only difference being that the mouse gDMR contains CTCF binding sites in which the insulator function causes maternal expression. The lack of a CTCF binding site at the human *GRB10* gDMR is believed to be the explanation for the biallelic expression in most tissues (Hikichi et al., 2003; Plasschaert and Bartolomei, 2015).

Figure 4.2

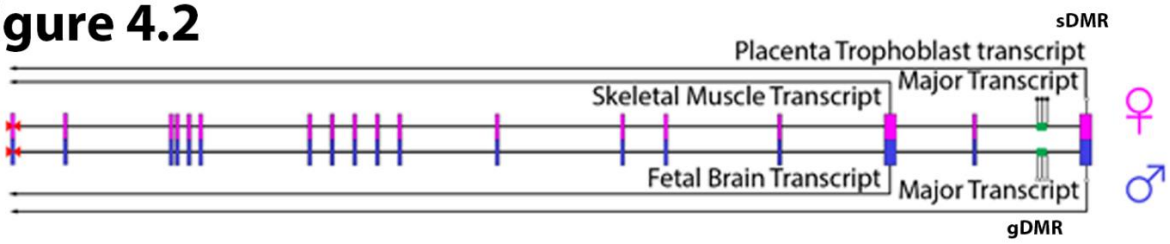


Figure 4.2: *GRB10* is an imprinted gene that is expressed in an isoformic and tissue-specific manner. The major transcript is biallelically expressed from the first transcription start site (TSS)in almost all tissues except in placenta trophoblast where a maternally expressed transcript is found. Further downstream there is a secondary TSS. Transcripts expressed from this TSS are only expressed in a monoallelic fashion with maternal expression found in the skeletal muscle and paternal expression found in the fetal brain. Methylation is consistent throughout all tissues suggesting that transcription factors rather than methylation may control the expression of GRB10. However, methylation in the adult brain is different with the somatic DMR at the first TSS containing methylation consistent to that of a DMR. Key unchanged from figure 1.5.

4.1.4 Methylation controlled expression of NLRP2

NLRP2 belongs to the nucleotide-binding, leucine-rich repeat-containing proteins (NLR) protein family, which is a member of the inflammasome and was the first recorded imprinted gene involved with inflammation (Agostini et al., 2004; Bruey et al., 2004). It was first discovered to have preferential maternal expression in the placenta and allelic skewing in kidney and fetal heart (Bjornsson et al., 2008). Unlike in human, *Nlrp2* in mouse does not appear to be imprinted, however, it plays an important role in early embryogenesis after oocyte activation, as female Nlrp2 null mice were likely to show early embryonic death of their offspring, while surviving offspring had aberrant methylation at imprints, most likely due to the abnormal localisation of DNMT1 in the oocyte due to loss of the SCMC complex (Kuchmiy et al., 2016; Mahadevan et al., 2017). Due to this, human NLRP2 is suspected to be associated with the SCMC complex as well, supported by its high expression in oocytes and preimplantation embryos. Furthermore, a mother with a germline homozygous frameshift NLRP2 mutation resulted in familial BWS affecting 2/3 of her offspring (Meyer et al., 2009; Monk et al., 2017). In somatic tissue such as fibroblasts and ES cells, *NLRP2* has a low level of expression though in induced pluripotent stem cells (iPSC) *NLRP2* is highly expressed suggesting it may have an important role in reprogramming (Cai et al., 2015). A study from Jefferies *et al* showed iPSCs to have a random imbalanced allelic expression between two clones with one clone showing biallelic expression and monoallelic expression in the other clone. However, interestingly unlike the other randomly imbalanced allelic genes seen in this study, *NLRP2* expression became monoallelic in the clone that was originally biallelic after neuralization (Jeffries et al., 2016). Evidence supporting the imprinted status of NLRP2 is the monoallelic expression noted by Bjornsson et al (2008), matched with the tissue-specific DMR described by Woodfine et al (2011) (Figure 4.3). While the sDMR was not clearly investigated by Woodfine *et al* (2011), the tissue-specific nature of DMR would suggest somatic establishment. Therefore, there may be a gDMR present within or near the gene body of NLRP2, with this in mind methylation at the promoter has been found to be inversely correlated with the expression of NLRP2 and for this reason, it would be a strong candidate region for further investigations to explain the monoallelic/imprinted expression of NLRP2 (Thürmann et al., 2018).

Figure 4.3

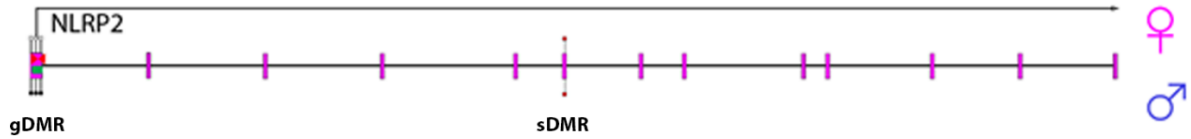


Figure 4.3: The imprinting of *NLRP2* has not been well characterised, the expression of *NLRP2* however, is negatively correlated with the level of methylation at the TSS, for this reason, it has been labelled and will be investigated as a possible gDMR in this study, a DMR has been described downstream, however, its establishment and parent of origin is unknown, though as it is tissue-specific it will be investigated as an sDMR candidate.

4.1.5 Uniparental ESCs

The generation of androgenetic and parthenogenetic embryonic stem cells furthered the ability to understand the mechanism behind genomic imprinting. Mouse parthenogenetic ESCs (phESC) were the first uniparental cell line generated firstly by activation of an oocyte without fertilisation (parthenogenesis), later that same year another lab generated mouse phESCs by removal of the male genome after fertilisation (gynogenesis) (Kaufman et al., 1983; Surani and Barton, 1983). It wasn't until a few years later that the mouse androgenetic ESC (aESC) cell line was generated, by androgenesis in which an oocyte void of its maternal genome has a paternal genome introduced (Mann et al., 1990). In 2007 two labs used the parthenogenesis technique to create human phESCs from parthenogenetic blastocysts which genetically matched their donor oocyte and showed methylation patterns at the imprints that were of parthenogenetic origin (Kim et al., 2007; Revazova et al., 2007). It wasn't until more recently that a human aESC cell was generated, they were observed to have maintained a certain level of sperm methylation at the imprints (Ding et al., 2015). A recent study by Sagi *et al* (2019) showed the power behind human uniparental ESCs, by utilising methylation and expression array technologies they were able to identify parental specific methylation and expression of known imprints and even identified several possible novel imprints and even suggests the possibility of additional imprinted genes still to be found that have a more complex nature.

4.1.6 In vitro Models investigating DNMTs in imprinting

The use of mouse ESC (mESC) is a well-recognised approach for investigating the effect of the loss of important DNMTs (Li et al., 1992). DNMT KOs are tolerable with numerous knockouts being carried out from 1 ESC to triple KO ESCs (TKO) (Tsumura et al., 2006). The Walsh lab has previously shown that imprinting could be recovered in mouse *Dnmt3a2* KO ESCs upon rescue with *Dnmt3a2* but was unable to recover imprinting in *Dnmt1* KO mESCs with the subsequent rescue of *Dnmt1*. This has been beneficial to increase the understanding of *Dnmts* in mouse, however, the KO of DNMT1 could not be replicated in human ESCs with the KO appearing to be lethal (Liao et al., 2015). A DNMT1 hypomorphic mutation could, however, be tolerated in a colorectal carcinoma line known as HCT116 (Rhee et al., 2000), but this is not suitable for imprinting studies due to the fact that it is aneuploid. Instead, the Walsh lab used TERT-immortalised fibroblasts known as *hTERT*-1604s to generate isogenic human cell lines with intermediate knockdown of DNMT1 (labelled as d8 and d10) and a more severe knockdown (labelled d16 -Figure 4.4 A and B) (O'Neill et al., 2018). The *hTERT*-1604 cell line was immortalised by overexpression of the human telomerase gene ensuring chromosomal integrity making it a highly suitable model for the study of methylation especially for imprinting (Ouellette, 2000). A study carried out by the Walsh lab showed robust demethylation effects in *hTERT*-1604 cells after a *DNMT1* knockdown (KD) (Loughery et al., 2011). The methylation microarrays showed that d8 and d10 had a similar loss of methylation while d16 had a more extreme loss of methylation when compared to the WT (Figure 4.4C). However, a more in-depth analysis revealed that 1/3 of the probes lost methylation as expected but surprisingly the remaining 1/3 of probes were observed with hypermethylation with targets shared among the 3 clonal DNMT1 knockdowns (Figure 4.4D). Using this DNTM1 depleted cell line our lab has recently highlighted sensitive genes that have an interesting interplay with polycomb repressive marks appearing to block remethylation. As methylation plays an important role in the regulation of imprints, this model is an excellent opportunity to investigate the long-term depletion of DNMT1 in a human differentiated cell line.

Figure 4.4

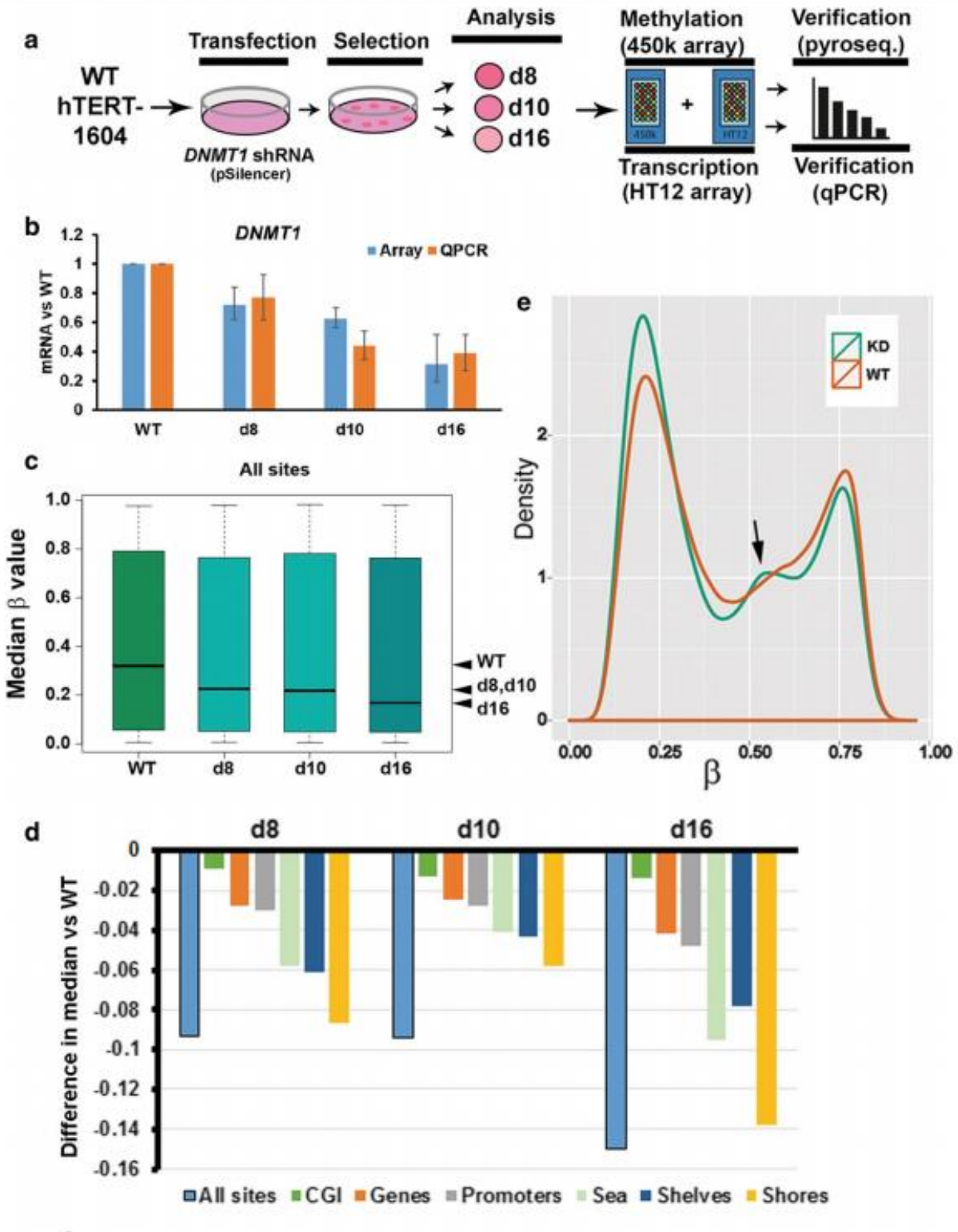


Figure 4.4: Establishment of DNMT1 depleted hTERT-1604s and changes of methylation. (A) Shows the experimental approach taken to generate and analyse the DNMT1-depleted hTERT-1604s. Transfection was carried out using an shRNA targeting DNMT1 and then cells grown in selective media until colonies appeared. Three separate clonally-derived cell lines were expanded (d8, d10 and d16) and further analysed with methylation and transcription arrays and confirmed with wet-lab techniques. (B) Confirmation of DNMT1 depletion in the d8, d10 and d16 cell lines using the HT12 array and qPCR: error

bars represent 95% confidence intervals around the median and standard error of the mean (SEM), respectively. (C) Overall changes of methylation observed across all probes on the Infinium human 450k array: a βvalue of 1 equates to 100% methylation. Median values are indicated by the line and whiskers represent interquartile range. The positions of the medians are also indicated at right (arrowheads). (D) The difference in median of DNMT1 depleted cells lines vs WT at different CG regions (E) The number of sites showing loss and gain of methylation in d8, d10 and d16 per no. of probes $x10^4$. Figure derived from O'Neill et al (2018)

4.2 Hypothesis and aim

We hypothesise that the promoter region of *NLRP2* controls the monoallelic expression of the protein. Secondly, as DNMT1 provides an important role in the maintenance of genomic imprinted regions a long-term depletion will result in abrogation of the imprinted marks.

The main aims of this study were to:

- Investigate the suspected imprinted status of *NLRP2* by using publicly available parental specific ESC datasets. The imprinting of *NLRP2* will be further investigated using the DNMT1 depleted chromosomally stable hTERT-1604 cell lines.
- Secondly, we aim to use the 450k array data generated by the Walsh lab, to investigate the impact of the long-term depletion of DNMT1 on the imprinted gametic and somatic DMRs in the chromosomally stable human fibroblastic cell model hTERT-1604.
- Known and well described imprinted regions that show large changes of methylation will also be investigated with wet lab techniques to fully understand any mechanistic changes occurring due to the long-term depletion of DNMT1 in a human chromosomally stable model.

4.3 Materials and Methods

Methylation analysis of the 450k and pyrosequencing and expression analysis by qPCR was carried out as described in chapter 2. Uniparental human ESCs publicly available methylation and expression array datasets were analysed (GSE114679). Publicly available ChIP-SEQ ZFP57 and ZFP445 in hESC datasets (GSE57989) were utilised to generate tracks on UCSC. 850K methylation datasets were downloaded and processed as described in chapter 2. GSE114679 came from a study by Sagi *et al* (2019), using single-cell genome-wide analysis for methylation

(850k beadchip array) and RNA-seq on androgenetic and parthenogenetic stem cells. This allowed Sagi *et al* (2019) to investigate methylation bias and distinct imprinting signatures. The second dataset GSE57989 was from a study carried out by Takahashi *et al* (2019), this investigation contributed to the literature by furthering our understanding of ZFP binding and imprinting in both human and mouse models.

4.3.1 UseGalaxy RNA-seq analysis

RNA-seq data was analysed using USEgalaxy, the data was uploaded to usegalaxy from the European nucleotide archive (ENA) in the compressed fastqsanger.gz format. First, the data was subjected to quality control which was tested by the FastQC (v 0.72+galaxy1) and carried out by the Cutadapt (v1.16.6) tool to trim low-quality sequences. The data was then mapped to the reference genome hg19 using the HISAT2 (v2.1.0+galaxy5) tool providing a BAM output file which was then processed with the featureCounts (v1.6.3) tool providing a table containing Entrez gene ID with its respective count. The Entrez Gene ID was then used to annotate the genes with Ensemble ID, Entrez ID, gene symbol and gene name using the tool annotatemyids (v3.7.0). Differential expression was carried out with limma-voom using the limma (v3.38.3+galaxy3) tool.

4.3.2 UseGalaxy ChIP-seq analysis

ChIP-seq data was analysed using USEgalaxy, the data was uploaded to usegalaxy from the European nucleotide archive (ENA) in the compressed fastqsanger.gz format. First, the data was subjected to quality control which was tested by the FastQC (v 0.72+galaxy1) and then manipulated to remove low quality reads with the Filter by quality (v1.0.2) and FASTQ Quality Trimmer (v1.1.1) tools. The data was then mapped to the reference genome hg19 using the Map with BWA (v0.7.17.4) tool providing a BAM output file. Data was then subjected to post-map processing which involved the removal of duplicates RmDup tool (v2.0.1). The bamCoverage (v3.3.0.0.0) tool was then used to generates a coverage bigWig file from the BAM dataset allowing for the data to be visualised on UCSC.

Table 4.1- RT-qPCR primers used in this chapter

Gene	Primer	Oligo sequence (5’-3’)
GRB10	FWD	GCACGAAGACAACCAGGTG
	REV	GAACGATCACTGCCTTACCC
SNRPN	FWD	CACCAAGAGGTGGTTAAAGC
	REV	GATTGCTGTTCCACCAAATCC
NLRP2	FWD	CGGCCGAGAGAGAAGCCTTATTAG
	REV	CAACACCTGGCCCTACTCGC

Table 4.2- Pyrosequencing primers used in this chapter

Gene	Primer	Oligo sequence (5’-3’)
GRB10 gDMR	FWD	GGGAGAAAGAGGTTTTTA
	REV (Biotin)	AAATCTAAACATCC
	SEQ	ATTTAAAAAATAAATAAATCTAAACATCC
GRB10 sDMR	FWD	AGTGAAAGGATAAATTGGATT
	REV (Biotin)	CCAAATTTCCTTTTCCCAAACC

	SEQ	GAAAGGATAAATTGGATTTG
SNRPN gDMR	Qiagen Assay	Hs_SNURF/SNRPN_01_PM PyroMark CpG assay
ATP10a sDMR	FWD	GGAGGTATAGAGAGGTTTTG
	REV (Biotin)	AAATACTTAAAAAACAACTCCACTCA
	SEQ	AAATACTTAAAAAACAACTCCACTCA
NLRP2 gDMR	FWD	GGGATTGGTTTGAATTGTAG
	REV (Biotin)	CCTTAAATAACATCACCCTATTCAACA
	SEQ	ATTGGTTTGAATTGTAGGA
NLRP2 sDMR	FWD	GATGGAGGATTATAGGTGGAGAT
	REV (Biotin)	TCCCAACACCTAACCCTACT
	SEQ	GGATTATAGGTGGAGATT

4.4 Results

4.4.1 Imprinting of NLRP2

Imprinting of *NLRP2* is not clear and not well investigated in the literature, initial analysis of the *NLRP2* promoter region reveals a methylation pattern indicative of maternal methylation bias in genome-wide UPD patients (Figure 4.5). Further to this figure, 4.5B showed that the methylation variability at the NLRP2 promoter region is higher than that of what is seen at other imprinted DMRs and confirms that the IG-DMR is covered in the 850K array. To further understand the imprinting behind this region and confirm if it is imprinted, a 450k and an RNA-seq dataset for uniparental ESCs were utilised. To ensure that the methylation array shows normal methylation patterns the methylation was examined in the aESCs and phESCs and compared to WT ESCs at the *H19* (paternally methylated) and *SNURF* (maternally methylated) DMRs. This showed hypermethylation in the aESCs and hypomethylation in phESCs at the paternally methylated *H19* DMR as expected (Figure 4.6A). The methylation at the maternally methylated DMR showed the opposite response with hypomethylation in aESCs and hypermethylation in the phESC at the *SNURF* DMR as also expected (Figure 4.6B). To ensure that these changes were unique to the imprinting DMRs the methylation was examined at the normally hypermethylated and hypomethylated CGIs at the non-imprinted genes *DAZL* and *MTHFR* promoter regions respectively, in which no change of methylation was observed between the WT and uniparental ESCs (Figure 4.6C & D).

Figure 4.5

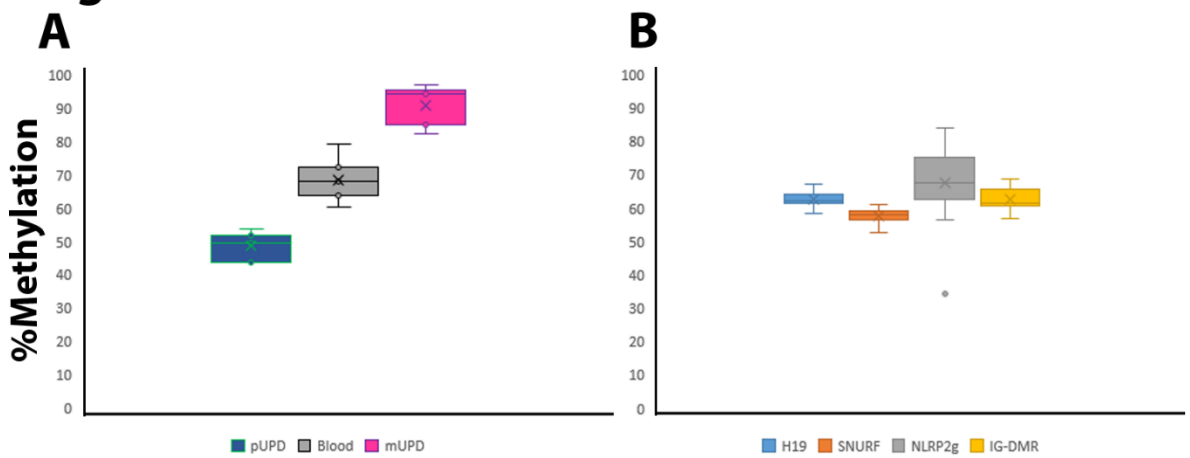


Figure 4.5: Methylation bias at NLRP2 using the beadchip array and variability of methylation at a subset of imprinted DMRs. NLRP2 promoter methylation bias in pUPD (n=1) and mUPD (n=1) in comparison to control blood (n=5). **(B)** Investigating the DMR variability between shows that the while the *IG-DMR* contains low probe coverage in the 850k epic array it still presents with tight variability across 23 healthy Saliva samples, similar to what is seen for the *H19* and *SNURF* DMRs. Across 23 saliva samples, the suspected *NLRP2* DMR presents with large variability which may suggest that this region is not imprinted in saliva. The box represents the interquartile range (IQR), the whiskers are IQR*1.5; the X marks the mean and a line represents the median; circles are outlier probes. MWU significance in comparison to blood demonstrated by *p < 0.05; **p < 0.01; ***p < 0.001.

Figure 4.6

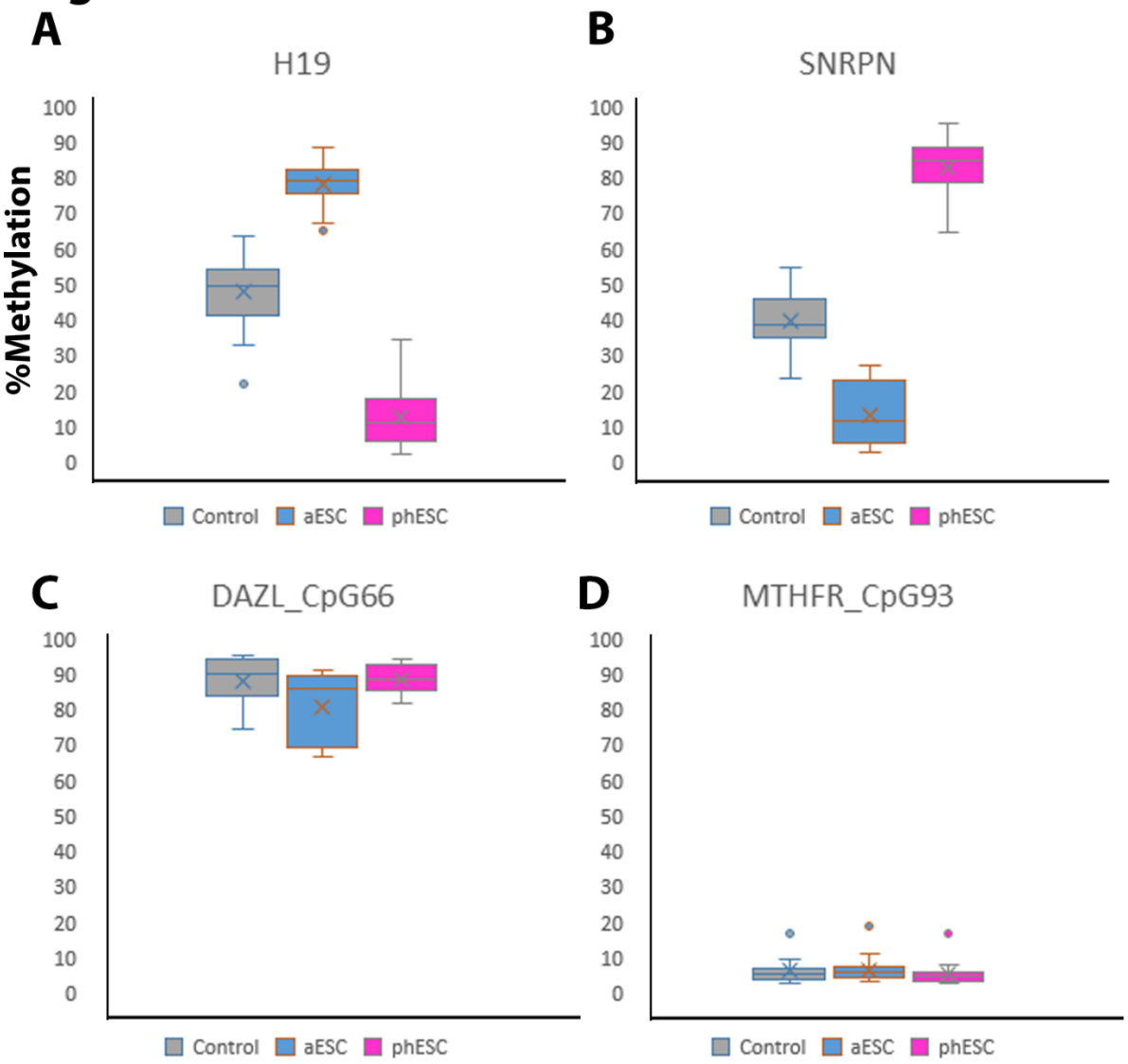


Figure 4.6: Establishing methylation level at imprinting DMRs in aESCs and phESCs. Utilising 450k array data for uniparental ESCs, it is possible to see distinct changes of methylation at the imprinting DMRs near the (A) *H19,* (B) *KvDMR* and (C) *SNURF* loci in comparison to the control WT ESCs. To ensure that the changes of methylation between the uniparental ESC samples and WT ESCs were unique to the imprinting DMRs methylation was examined at the heavily methylated CGI at the (D) *DAZL* promoter region and the lowly methylated CGI at the (E) *MTHFR* promoter region. X; Mean, Line; Median, Box; interquartile range (IQR), Whiskers; IQR*1.5, Circles; probe outliers. aESC, androgenetic ESCs; phESC, parthenogenetic ESCs.

Next, it was important to confirm the DMR at the promoter region, to achieve this, each CpG site was examined from 5' end to the 3' end of *NLRP2* to fully investigate the imprinting at this region. This showed clear hypomethylation at the first 8 CG sites which represent the promoter region which was covered by the array in the phESCs and hypermethylation at the same CG sites at the aESCs (Figure 4.7). The remaining of the CG sites covered by probes covering the gene body show hypermethylation (Figure 4.7). Suggesting that *NLRP2* has a paternal methylation bias similar to the *H19* DMR. While both parental UPD patients and parental ESCs models showed a methylation bias, confusingly they are opposite in which the UPD patients show maternal methylation and the ESC's show paternal methylation.

Figure 4.7

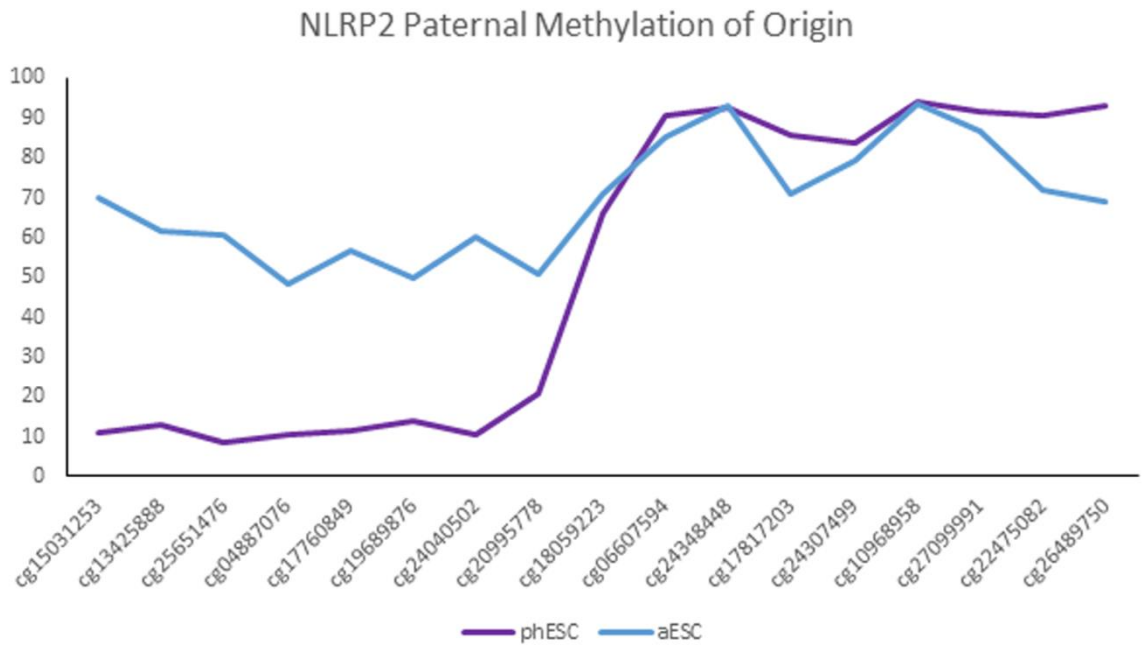


Figure 4.7: Methylation at each probe covering *NLRP2* in phESC and aESCs. Methylation at the probes in the 450k array, showing methylation paternal methylation bias in the aESC compared to the phESC at the promoter region (First 8 CG sites), then hypermethylation in both the aESC and phESCs at the probes covering the gene body region of *NLRP2*. aESC, androgenetic ESCs; phESC, parthenogenetic ESCs.

After confirming the methylation bias in a second model for imprinting, RNA-seq analysis was utilised to investigate parental specific bias. Firstly, parental specific expression was examined in known maternally and paternally expressed genes. The maternally expressed genes *MEG3*, *H19* and *CDKN1C*, were observed with upregulation in the phESCs in comparison to the WT ESC controls, while downregulation was observed in the aESCs in comparison to the WT ESCs controls (Figure 4.8 A & B). Contrary the paternally expressed genes *PEG3*, *IGF2* and *SNRPN* were observed to have downregulated expression in the phESCs in comparison to the WT ESCs, while in the phESCs they are observed to have upregulation in aESCs (Figure 4.8 A & B). To ensure that the methylation bias observed at the imprinted DMRs were controlling the parental specific expression in these genes, *DAZL* and *MTHFR* genes were examined which were observed to have hypermethylation and hypomethylation at the CGIs at their promoter region respectively, both these genes showed similar expression trends in both the phESC and aESC in comparison to the WT ESCs confirming that the parental specific expression seen in these ESCs is unique to the Imprinted genes (Figure 4.8 A & B). As the methylation bias at the imprinted DMR methylation and gene expression is confirmed to followed expected patterns, the NLRP2 gene expression was then examined which showed upregulation in the phESCs in comparison to the aESCs, while in the aESCs it was observed to be downregulated in comparison to the WT ESCs (Figure 4.8 A & B). Taking this expression in account with the methylation data, *NLRP2* appears to be expressed from the unmethylated maternal allele in these cell models. As previously mentioned the androgenetic and parthenogenetic ESC's resemble the genome of their respective gametes and therefore would suggest that this region should be treated as a gDMR for the downstream analyses.

Figure 4.8

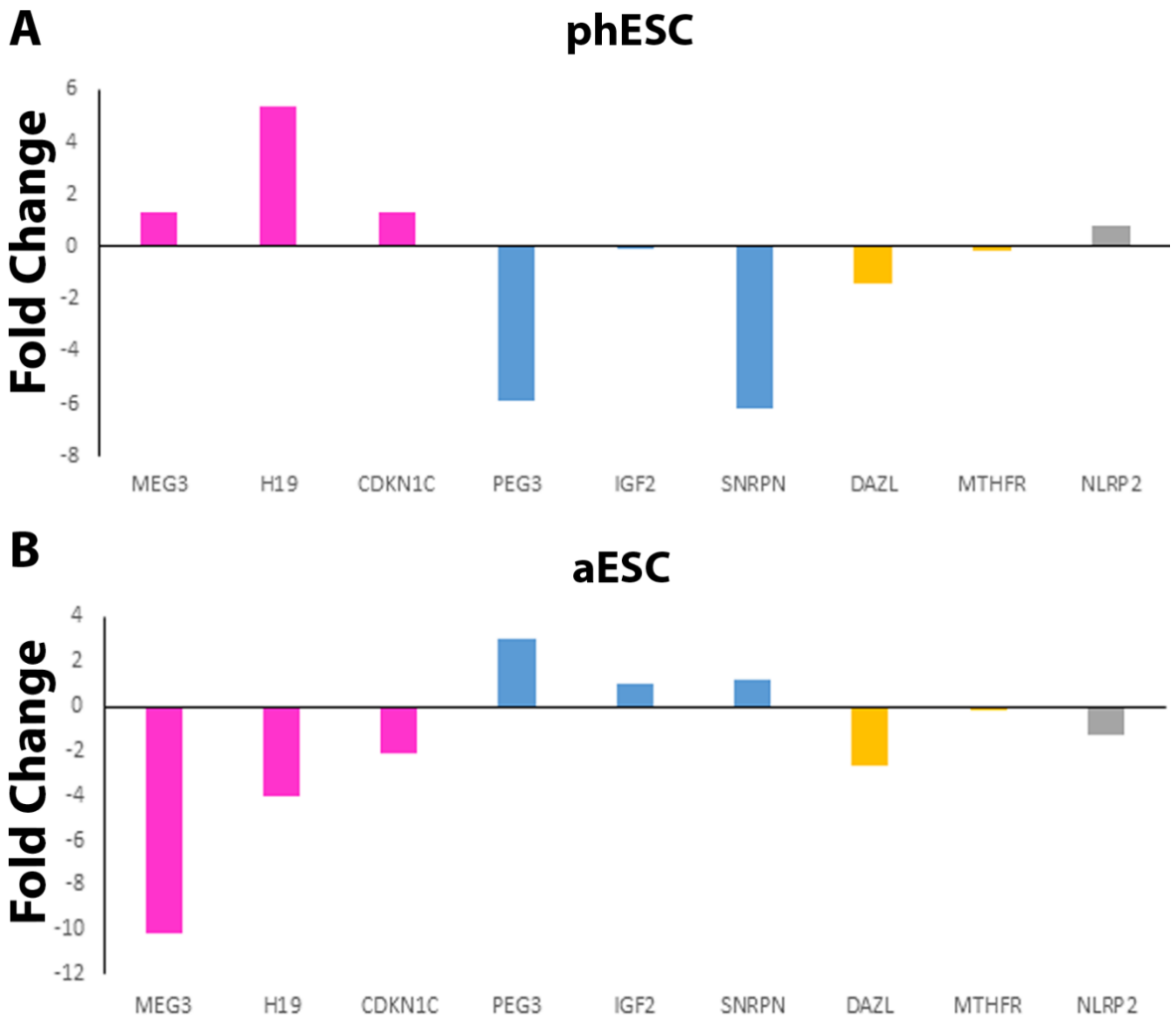


Figure 4.8: Single-cell RNA-seq expression analysis of a subset of imprinted genes androgenetic and parthenogenetic embryonic stem cells. RNA-seq analysis showed the expected parental specific expression for the maternally expressed genes (*MEG3*, *H19* and *CDKN1C*) and paternally expressed genes (*PEG3*, *IGF2* and *SNRPN*) in both **(A)** parthenogenetic ESCs (PhESCs)and **(B)** androgenetic ESCs (aESCs) vs WT control ESCs. NLRP2 appeared to have a biased expression increased expression in the phESCs vs the WT while aESC were observed with decreased expression vs WT control ESCs.

To finish the investigation at the possible imprinting mechanism at the *NLRP2* imprinted gene, histone modifications, ZFP57 and ZFP445 binding, CTCF binding and repetitive elements were examined using tracks on UCSC. Unexpectedly using all the histone modifications from the ENCODE project there were no histone modifications found at the *NLRP2* DMR (Figure 4.9). At the novel *NLRP2* DMR, there was no ZFP binding observed, but slightly upstream a small peak was observed for ZFP57 binding (Figure 4.9). Interestingly the whole region is full of SINE repetitive elements, except at the two different transcriptions start sites which both proceed LTR elements and correlate with the probes from the 450k array (Figure 4.9). While a clear methylation bias is observed at the promoter region, the imprinted status is still not clear due to the lack of other common epigenetic markers found at imprinted DMRs. However, as there is a clear allelic methylation bias, the region will be included in downstream studies in hopes that further investigations may aid in understanding what is occurring at this region.

Figure 4.9

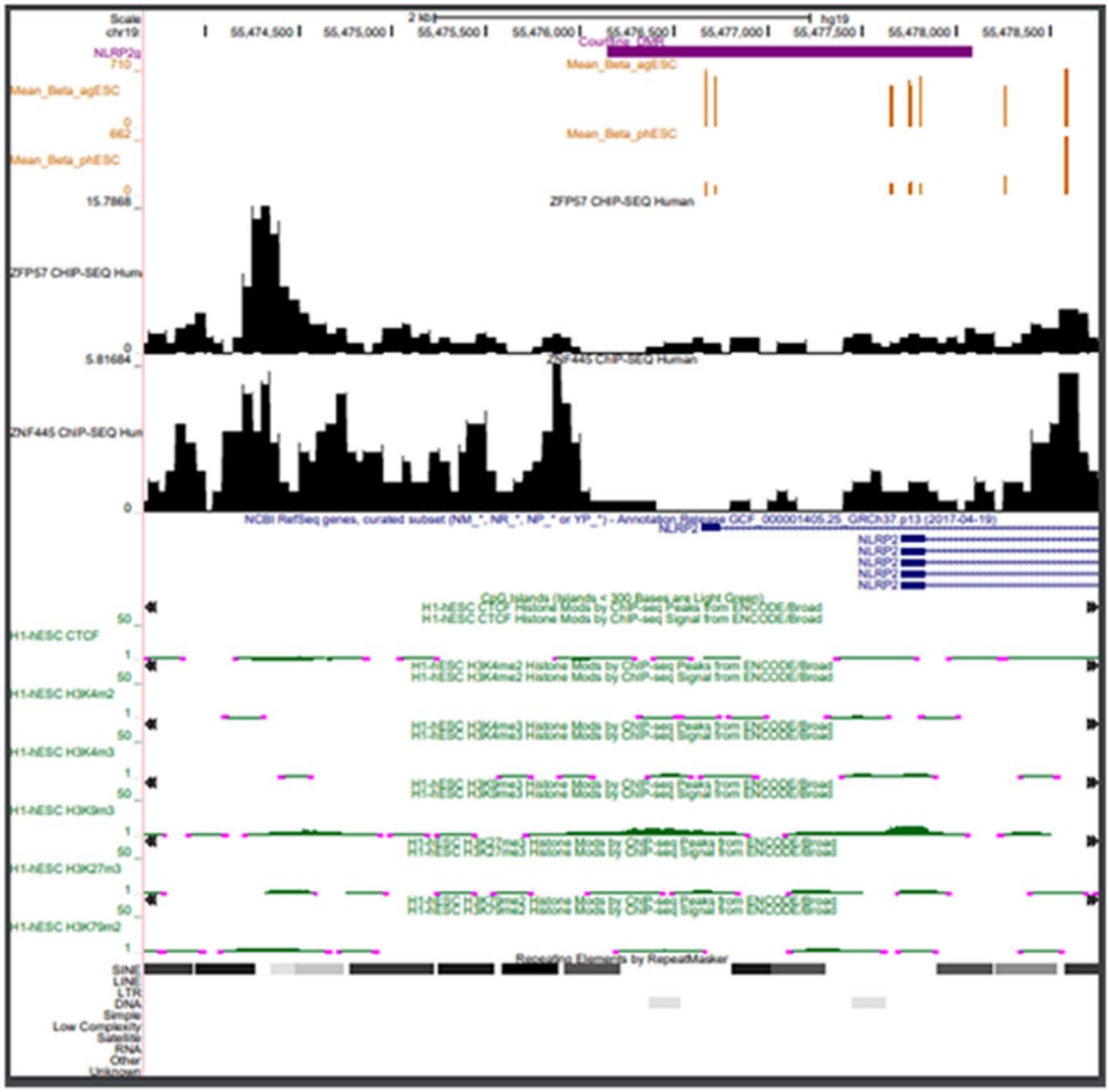


Figure 4.9: UCSC tracks describing epigenetic regulators near the NLRP2 promoter region. DMR locations are marked in purple, orange indicates beta methylation. NCBI RefSeq was used to show the location of the gene. Mean_Beta_agESC represents the paternal methylation out of 1000 while the Mean_Beta_phESC represents the maternal methylation out of 1000. Slight enrichment of ZFP57 is noticeable upstream from the novel gDMR and no enrichment of ZFP445 using hESC ChiP-SEQ data from (Takahashi et al., 2019). Using ENCODE tracks on UCSC no enrichment of histone marks could be observed. This region also contains many SINE repetitive elements nearby the transcription start sites (TSS) of *NLRP2*, though DNA repeats present just before the TSS according to the UCSC RepeaterMasker.

4.4.2 Aberrant methylation observed at imprinted genes in long term knockdown of DNMT1 in hTERT-1604s

The DNMT1 knockdown and WT cells had been analysed using the human Infinium 450k array (450k array), including subsequent visualisation of the data in the University of California, Santa Cruz (UCSC) genome browser, as tracks allowing. Through a bespoke GALAXY workflow developed in-house in the Walsh lab, the methylation values were analysed for the probes located at imprinted DMR coordinates as defined in Woodfine *et al* (2011) and Court *et al* (2014). The long-term depletion of DNMT1 resulted in widespread demethylation at most imprinted DMRs. However, a general trend is observed in which LOM is observed at the gDMRs (Figure 10A), while a number of DMRs characterised as somatic (sDMR) or unknown (uDMR) by Woodfine *et al* (2011) were seen to be gaining methylation as seen in Figure 4.10B and Figure 4.10C respectively. These unexpected gains of methylation provided interesting target regions for further analysis. I chose three loci which showed this clear trend for further analysis:1) the *GRB10* loci as it shows a large change at both the gDMR and sDMR; 2) the *SNURF/SNRPN* DMR and the *ATP10a* promoter region as they show loss and gain of methylation respectively and represent a well-understood imprinting locus in human and finally 3) *NLRP2*, as it showed a loss of methylation at our candidate gDMR matched with a considerable gain of methylation in d10 and d16 at the sDMR. All gametic, somatic and undescribed DMRs from the Woodfine *et al* (2011) and Court *et al* (2014) studies were included in this study. However, this did not include any placental DMRs found from the Court *et al* (2014) study.

Figure 4.10

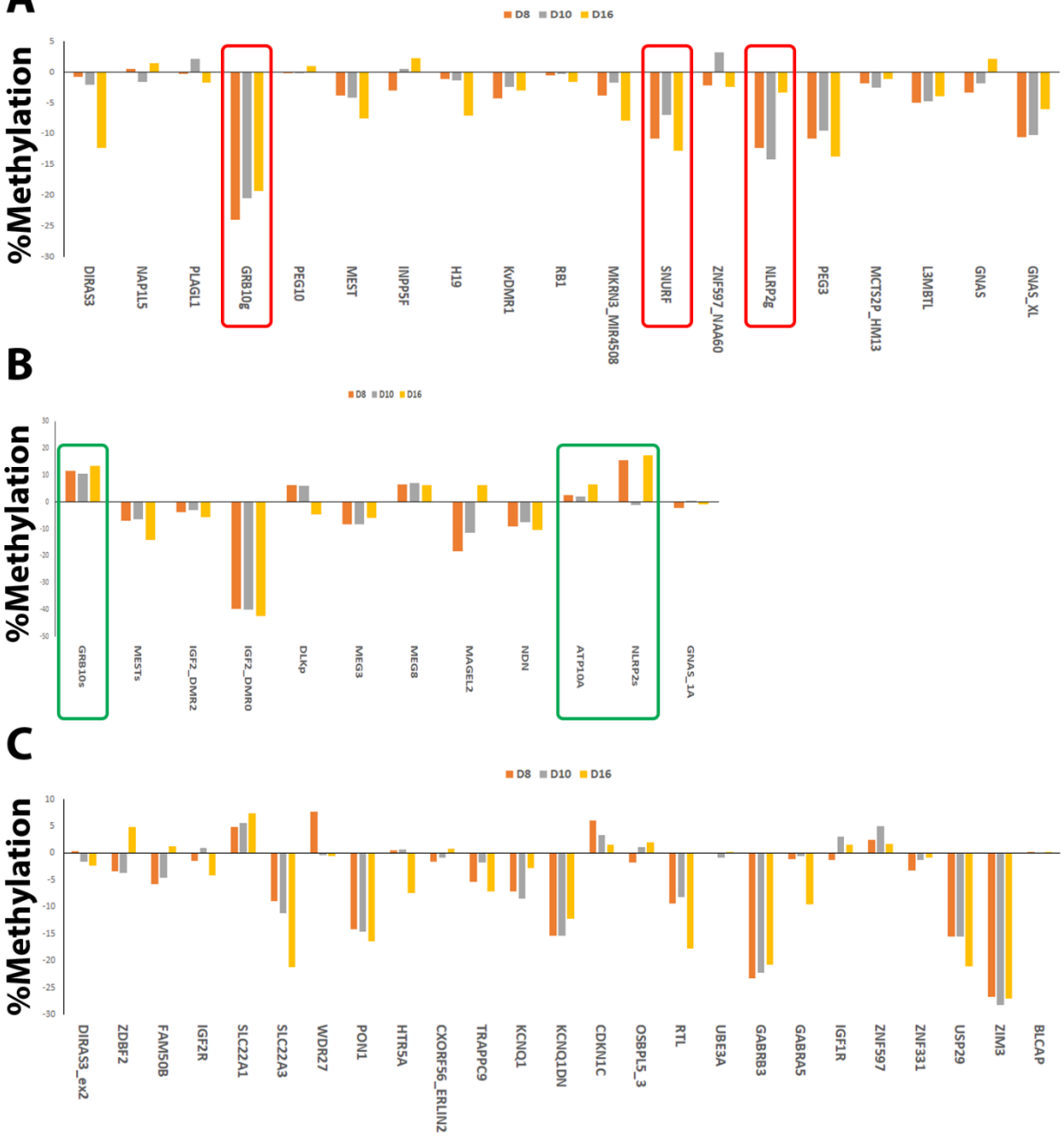


Figure 4.10: Changes of methylation (%) observed across DMRs found at/near imprinted loci in DNMT1 depleted hTER-1604s. The figure shows % changes of methylation in the clones D8, D10 AND D16 in comparison to WT at **(A)** Gametic DMRs, **(B)** Somatic DMRs and **(C)** unknown DMRs which have not been fully investigated for the establishment of the mark. Red boxes indicated gDMRs which lose methylation while green boxes indicate the nearby sDMRs of interest that show gains of methylation which will be further investigated in this study.

4.4.3 Changes of methylation at DMRs at the imprinted loci GRB10, SNRPN and NLRP2

The data from the DNMT1 knockdown cell line demonstrated an accumulation of hypomethylated probes at gDMRs and accumulation of hypermethylated probes at the sDMRs at the imprinted genes *GRB10*, *SNRPN* and *NLRP2* (Figure 4.11 A-D). Further analysis of the 450k array using the bespoke GALAXY workflow developed in-house was used to get mean methylation values at DMRs of interest, from this a switch-like mechanism of methylation between the gDMR and sDMR at the imprinted loci could be observed (Figure 4.12). Though all gDMRs investigated showed loss of methylation only the gDMR at the *GRB10* in the d8 cell line showed significant loss of methylation when compared to the WT. At the sDMR, *GRB10* and *NLRP2* show the biggest gains of methylation for all KD cells except interestingly d8. *ATP10a* gains of methylation are very low, though probe coverage for this region is very poor so it may not be an accurate representation. As probe coverage was limited at the DMRs this lowered the statistical power, therefore, methylation needed to be further confirmed with wet-lab techniques.

Figure 4.11

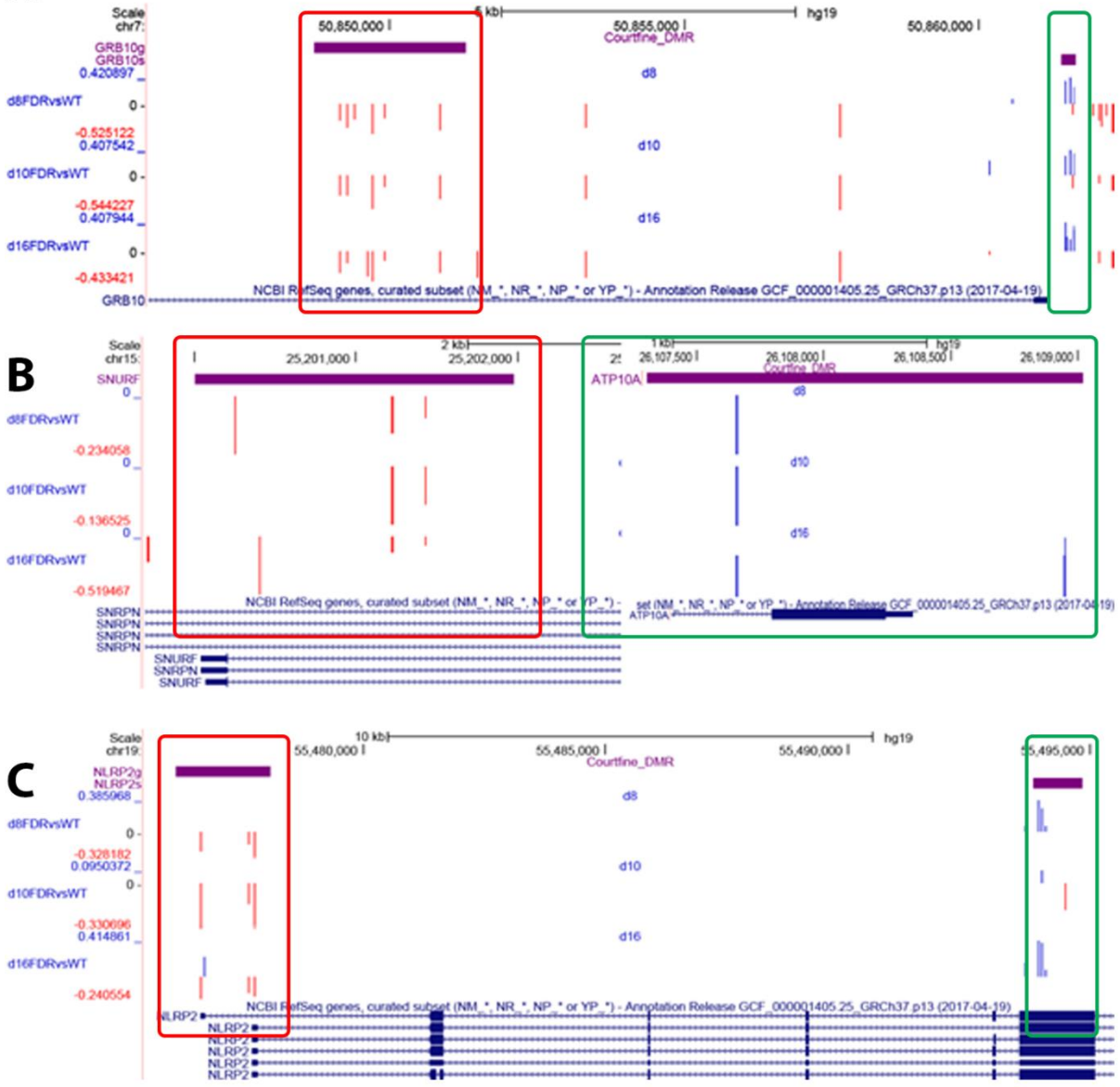


Figure 4.11: Long term knockdown by shRNA on hTERT-1604s led to an opposite methylation change at gDMR and sDMR at imprinted loci. DMR locations are marked in purple, Red indicates loss of methylation while blue indicates an increase of methylation. NCBI RefSeq was used to show the genomic location. FDR tracks compare the difference of the clonal cell lines d8, d10 and d16 to WT respectively. **(A)** The *GRB10* gDMR indicate a loss of methylation while the probes at the sDMR indicate an increase of methylation for all of the knockdown cell lines. **(B)** The probes located at the *SNRPN* gDMR indicate a loss of methylation for all of the knockdown cell lines. **(C)** The probes located at the *ATP10A* sDMR indicated a gain of methylation for all of the knockdown cell lines. **(D)** The *NLRP2* gDMR indicates a loss of methylation while the probes at the sDMR indicate an increase of methylation for all of the knockdown cell lines with the exception of d8 at the sDMR which appears to lack probe coverage in comparison to D8 and D10. Red boxes highlight the gametic DMRs losing methylation while the Green box is used to highlight somatic DMRs gaining methylation.

Figure 4.12

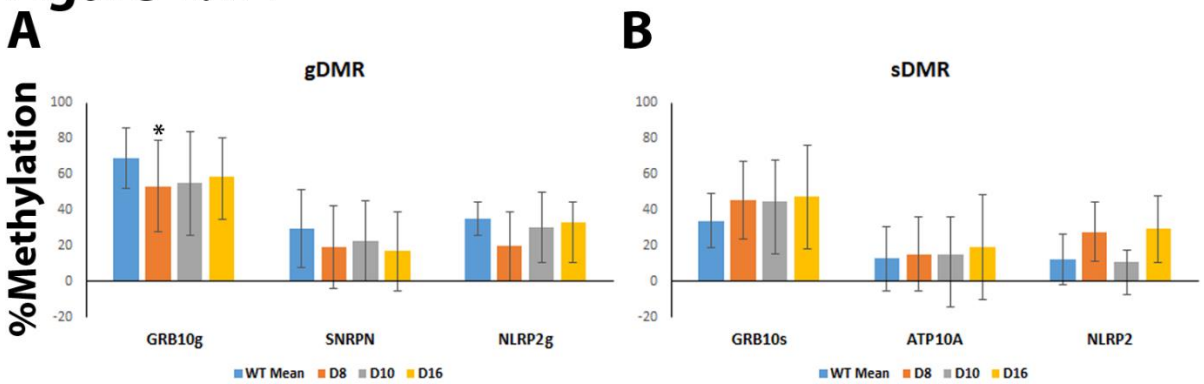


Figure 4.12: Average methylation on the 450k array was decreased at the gDMRs and increased at the sDMRs of imprinted genes in DNMT1 KD lines. CpG probe sites at DMR locations for select imprinted genes were averaged for WT hTERT-1604 and the clonal DNMT1 depleted cell lines (Blue), d8 (Orange), d10 (Grey) and d16 (yellow). **(A)** Average methylation at the gDMRs of the imprinted loci *GRB10*, *SNRPN* and *NLRP2* in WT, d8, d10 and d16. **(B)** Average methylation at the sDMRs of the imprinted loci *GRB10*, *ATP10A* and *NLRP2* in WT, d8, d10 and d16. 450K array was carried out for WT and each clonal cell line at passage number #22-#24 Error bars indicate standard deviation. Significance carried out with a Mann–Whitney U test vs WT with the following key, *p < 0.05; **p < 0.01; ***p < 0.001.

4.4.4 Confirming changes of methylation observed in the 450k array

To confirm the changes of methylation observed from the 450k array DNA was extracted from the WT and DNMT1 KD cell lines and underwent bisulphite conversion. Pyrosequencing was then utilised to confirm the changes of methylation observed in the 450k array. Pyrosequencing utilises sequencing by synthesis on bisulphite converted DNA, the methylation is quantified by the light given off as the region of interest is synthesized. Pyro assays were designed for the gDMR and sDMR at the *GRB10*, *NLRP2* and the sDMR at *ATP10a* while the *SNURF*/*SNRPN* assay was purchased from Qiagen. These pyro assays involved carrying out a PCR with bisulphite primers targeting the region of interest for *GRB10*, *SNRPN* and *NLRP2* as annotated on Figure 4.1, 4.2 and 4.3. A sequencing primer was then used to create a start point for the synthesis this was then sequenced and analysed using the Q24 pyrosequencer. At the gDMR of the loss of the imprinted gene, methylation was confirmed in each cell line compared to the WT. *SNRPN* showed significant loss of methylation in all cell lines as well as d8 at *GRB10* (Figure 4.13A). Further to this, the gains of methylation were also confirmed at each of the loci for each cell line. However, none appeared to be significant (Figure 4.13B). As methylation controls the expression at imprinted genes, I then aimed to

investigate the expression of these genes to fully understand the effect of this aberrant pattern observed

Figure 4.13

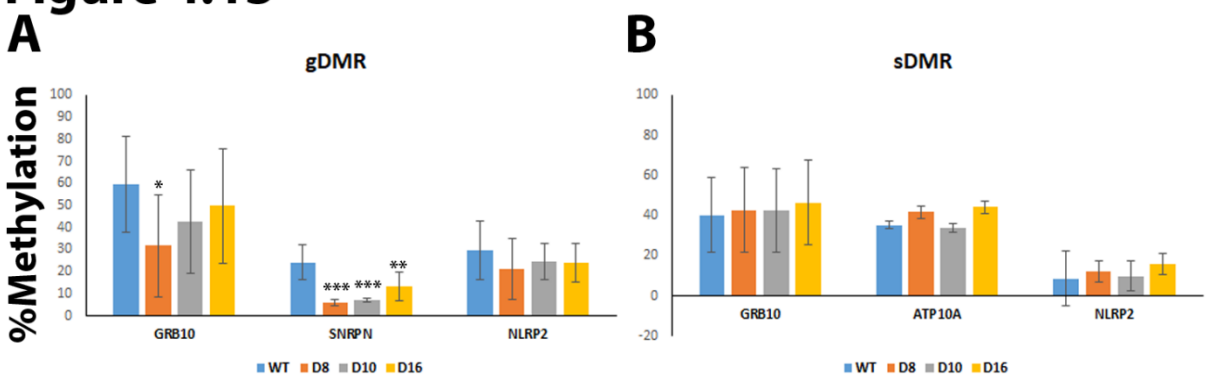


Figure 4.13: Confirmation of methylation changes by pyrosequencing at the gDMRs and sDMRs of imprinted genes in hTERT-1604 with long term shRNA DNMT1. CpG sites at DMR locations for select imprinted genes were averaged for the cell lines WT (Blue), d8 (Orange), d10 (Grey) and d16 (yellow). **(A)** Average methylation at the gDMRs *GRB10*, *SNRPN* and *NLRP2* WT, d8, d10 and d16. **(B)** Average methylation at the sDMRs *GRB10*, *SNRPN* and *NLRP2* WT, d8, d10 and d16. Pyrosequencing was ran in duplicate with two biological replicates. Pyrosequencing was carried out for WT and each clonal cell line at passage number #22-#24 Error bars indicate standard deviation. Significance carried out with a Mann–Whitney U test vs WT with the following key, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

4.4.5 Methylation pattern appears to have a functional effect on the expression of the imprinted genes

Using complementary DNA (cDNA) the expression was investigated using qPCR with SYBR green, SYBR green is a DNA dye which emits green light which is picked up by the light cycler as the cDNA is amplified. This is then normalised using the housekeeping gene *HPRT* the results are seen in Figure 4.14A-C. *GRB10* had significantly ($p < 0.001$) lost expression in all three knockdowns when compared to WT with a fold change of 0.22, 0.003 and 0.015 in d8, d10 and d16 respectively. *SNRPN* followed the same direction of *GRB10* with a significant ($p < 0.001$) fold change of 0.16, 0.15 and 0.28 in d8, d10 and d16 respectively. Finally, *NLRP2* appeared to have significantly increased expression, with a fold change of 7.77, 2.43 ($p < 0.01$) and 4.86 ($p < 0.05$) in d8, d10 and d16 respectively when compared to WT. The expression of *GRB10* and *NLRP2* was further investigated with RT PCR as seen in Figure 4.14C, this is a nonquantitative approach to gene expression but clearly shows *GRB10* losing expression in the KD cell lines

while *NLRP2* is gaining expression in the KD when compared to the WT and normalized by HPRT.

Figure 4.14

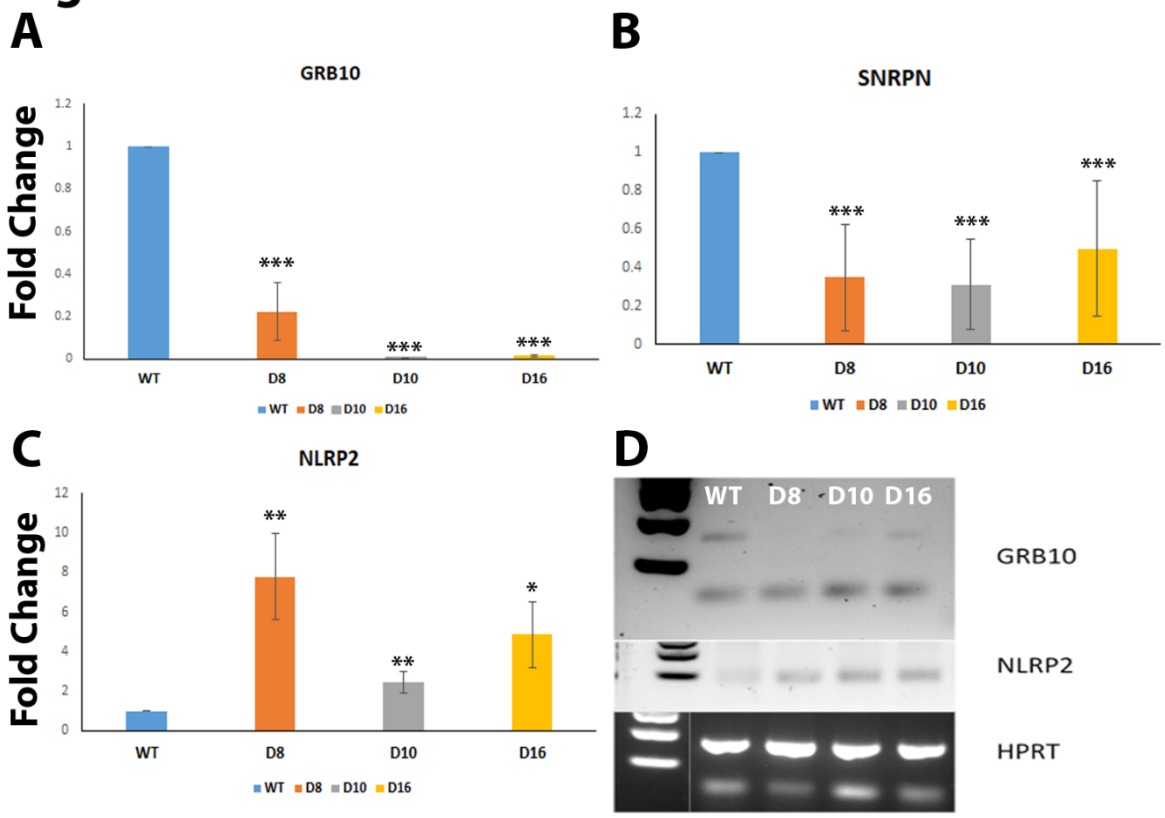


Figure 4.14: Aberrant methylation in long term shRNA knockdown of DNMT1 in hTERT-1604s correlates with aberrant expression of imprinted genes. (A-C) qPCR showing expression of the imprinted genes *GRB10*, *NLRP2* and *SNRPN* in d8, d10 and d16 when compared to WT, samples were normalised to HPRT. **(D)** RT PCR confirming the loss of expression in GRB10 and gain of expression in *NLRP2* with *HPRT* included for loading control. qPCR was ran in triplicate with two technical replicates and one biological replicate. WT and each clonal cell line were at passage number #22-#24 at point of transcription analysis. Error bars indicate standard deviation. Error bars indicate standard error of the mean. Significance carried out with a two-tailed T-test Vs WT, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

4.4.6 The gain of methylation at sDMR is unique to the long term knockdown of DNMT1 in the hTERT-1604s

As the interplay observed at the gDMR and sDMR of the imprinted genes investigated was unexpected, to determine whether or not the pattern observed was caused by the DNMT1 depletion itself, or was a result of subsequent changes in response to the knockdown which were selected for following the long-term culture of DNMT1-depleted clonal cell lines, a transient KD of DNMT1 in hTERT-1604s using siRNA was utilised. The siRNA is a double-stranded RNA that causes a knockdown of the targeted gene as long as the siRNA is being transfected into the cells. The siRNA has an overhang on the 3' end which activates the RNA interference which leads to the degradation of the targeted mRNA. The 450k array was utilised once more to measure the methylation then compared to the WT. Investigation of the data on UCSC demonstrated an accumulation of hypomethylated probes at gDMRs of *GRB10*, *SNRPN* and *NLRP2*. However, unlike the shRNA, the sDMR at these imprinted genes also demonstrated an accumulation of hypomethylated probes (Figure 4.15 A-D). This is further confirmed by averaging the probes in these regions as seen in Figure 4.16 A and B, which shows a slight loss of methylation however, the decrease was not significant. This data present here suggests that the interplay between the gDMR and sDMR of these imprinted genes observed in the shRNA KD is unique to the long term culture of the DNMT1-depleted cells. This can be seen globally for all imprinted DMRs in which the transient siRNA knockdown shows loss of methylation at almost every DMR (Figure 4.17).

Figure 4.15

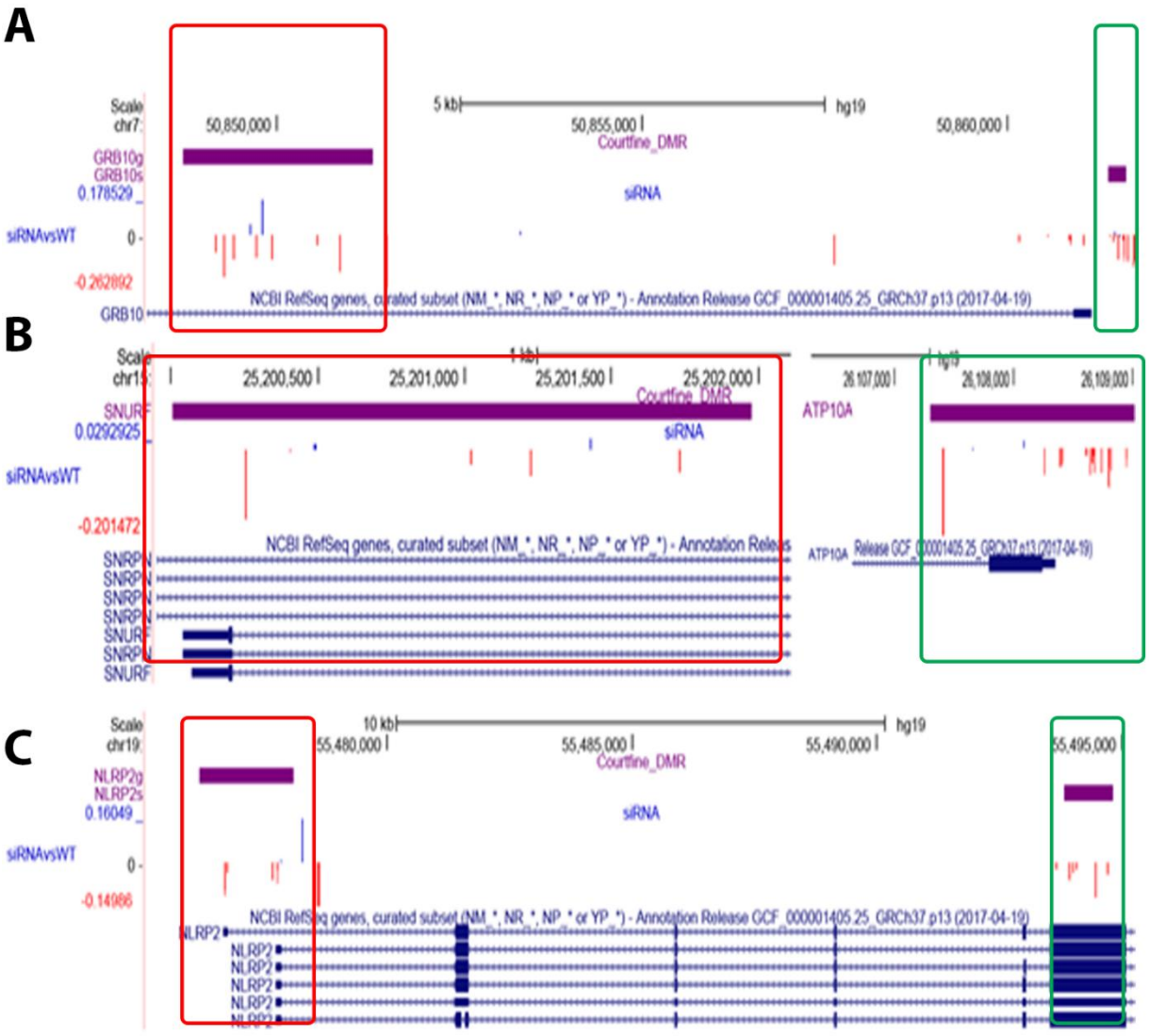


Figure 4.15: Transient knockdown by siRNA on hTERT-1604s led to a loss of methylation at both the gDMR and sDMR at imprinted loci. DMR locations are marked in purple, Red indicates loss of methylation at the gDMR while green indicates loss of methylation at the sDMR. NCBI RefSeq was used to show the location of the gene. FDR tracks compare the difference of siRNA DNMT1 depletion to WT respectively. Both the gDMR and sDMR show loss of methylation at the imprinted loci **(A)** *GRB10*, **(B)** *SNPRN/ATP10a* and **(C)** *NLRP2*.

Figure 4.16

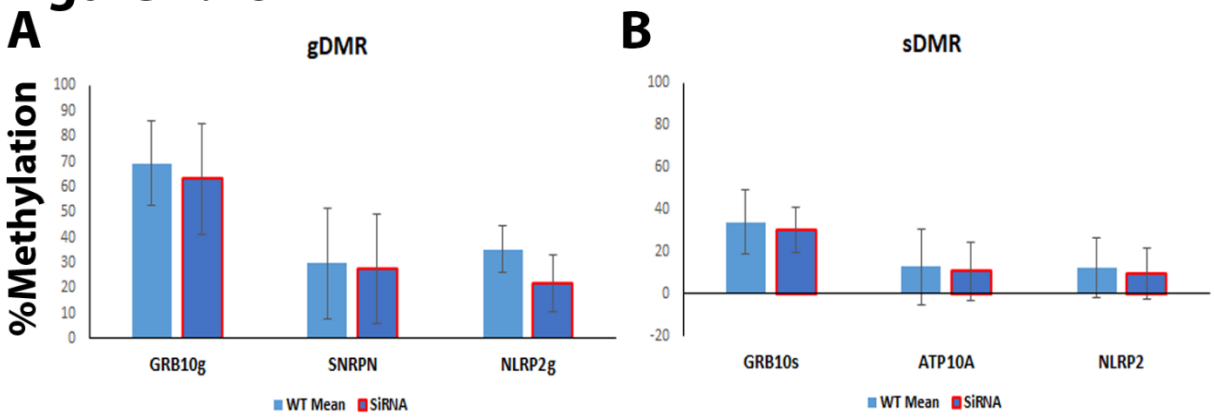


Figure 4.16: Loss of methylation at the gDMRs and sDMRs of imprinted genes in hTERT-1604 with transient knockdown of DNMT1. 450k array methylation levels at probes found with DMR locations for select imprinted genes were averaged for the WT and DNMT1 siRNA-depleted cell lines **(A)** Average methylation at the gDMRs of the imprinted loci *GRB10*, *SNRPN* and *NLRP2* in WT and siRNA cell lines. **(B)** Average changes of methylation at the sDMRs of the imprinted loci *GRB10*, *SNRPN* and *NLRP2* in WT and siRNA cell lines. Error bars indicate standard error of the mean. Significance carried out with a Mann–Whitney U test vs WT, however, no changes were significant.

Figure 4.17

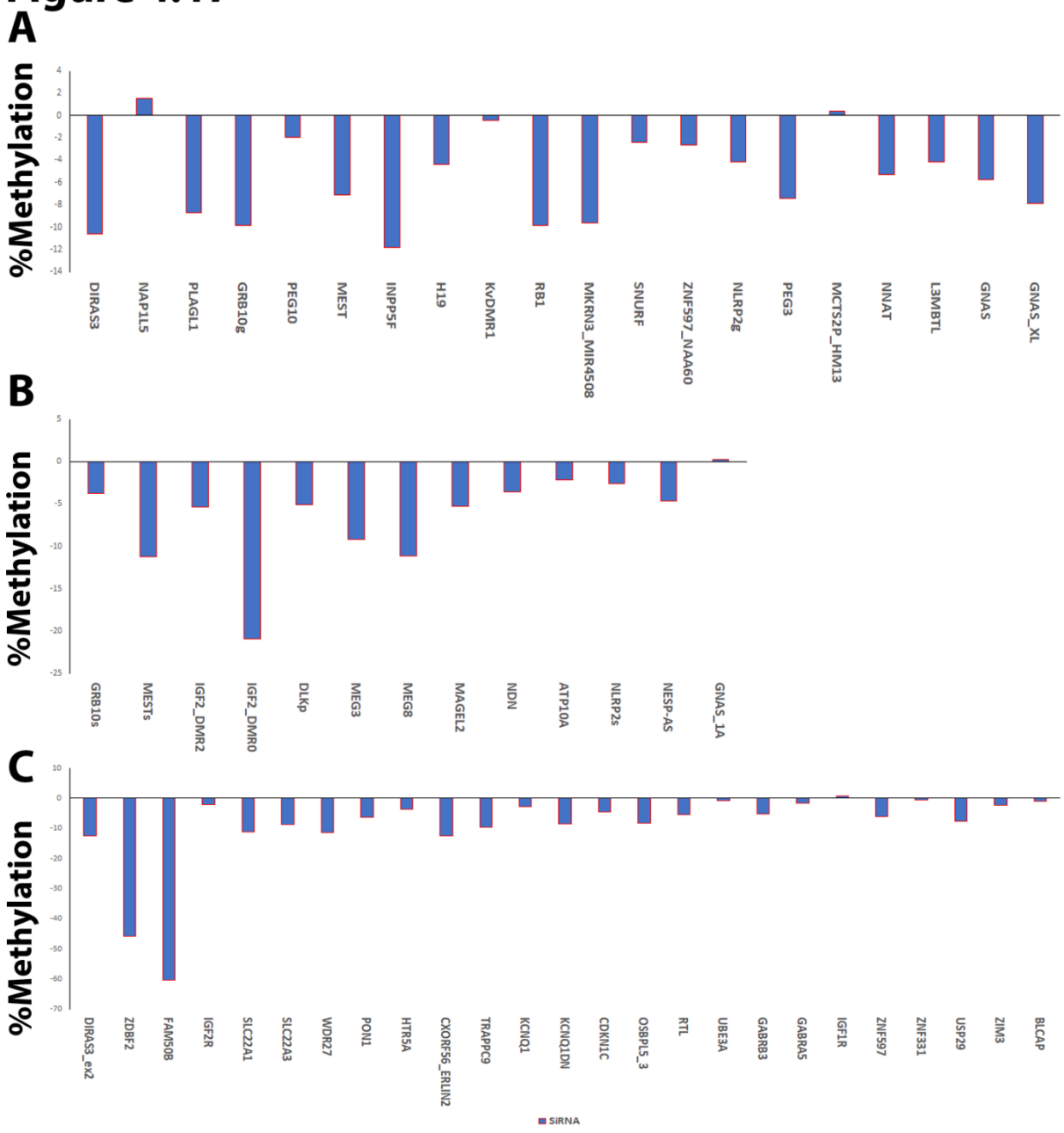


Figure 4.17: Changes of methylation (%) observed across DMRs found at/near imprinted loci in siRNA transient DNMT1 depletion vs WT. The figure shows % changes of methylation in siRNA transient knockdown of DNMT1 in comparison to WT at **(A)** Gametic DMRs, **(B)** Somatic DMRs and **(C)** DMRs which have not been fully investigated for the establishment of the mark.

4.5 Discussion

Methylation patterns at imprinted DMRs are important for normal regulation of imprinted genes and disruption at these DMRs can lead to imprinting disorders. One method to investigate the effect of aberrant methylation at imprinted DMRs is to deplete or knockout essential enzymes involved in de novo DNA methylation or maintenance. Aberrant expression of DNMT1 has been shown to affect global methylation in somatic tissue (Etoh et al., 2004). Human transient or long-term depletion of DNMT1 has been shown to cause loss of imprinting in human cancer cells transfected with siRNA and shRNA (Anwar et al., 2012; Min et al., 2017). However, these studies are limited due to cancer cell lines being chromosomally unstable as mentioned previously. To circumnavigate this the Walsh lab carried out a long-term knockdown of DNMT1 in a differentiated, chromosomally stable human fibroblast cell line and as expected global loss of methylation was observed. Imprinting marks established in the germline require DNMT1 to be maintained in somatic tissue, this explains the loss of methylation confirmed at the gDMRs in the DNMT1 depleted cell lines (Kurihara et al., 2008). However, a "meta-stable" switch was observed at some imprinted genes with the gDMR losing methylation and nearby sDMRs/uDMRs gaining methylation despite the loss of DNMT1. To ensure that the changes seen in the 450k array were accurately representing the methylation at the imprinting DMRs *GRB10*, *SNRPN* and *NLRP2* were analysed with pyrosequencing. Though most changes of methylation observed were statistically insignificant, the fact it was seen with two different methods would suggest that a meta-stable switch occurs in the DNMT1-depleted hTERT-1604s. One possible explanation for the hypermethylation could be an over-expression of *de novo* methyltransferases which have been known to have maintenance activity in somatic tissues (Lei et al., 1996). The majority of maintenance methylation in cultured adult human cancer cells is carried out by DNMT1 and DNMT3B (Rhee et al., 2002). While not shown in this thesis Investigation of the expression of DNMT3B in the DNMT1-depleted cell lines using the HT12 array showed no significant change of expression, suggesting that the gains of methylation are not caused by DNMT3B.

Imprinted gene expression is tightly regulated by methylation, for this reason, the next step was to investigate the expression of the 3 target imprinted genes to see how the aberrant methylation pattern affected the expression. Though the changes of methylation were not seen to be significant, the expression was reproducibly affected by the DNMT1 depletion. GRB10 is highly tissue-specific with the sDMR controlling the expression of the major transcript, while the gDMR is present on the TSS of the imprinted GRB10 transcripts that are expressed in a tissue-specific manner (Blagitko et al., 2000; Monk et al., 2009). GRB10 transcription analysis was carried out on different transcripts with primers binding to exons unique to each transcript, to investigate if the changes of methylation affected the transcripts differently, with loss of expression observed at each transcript investigated (data not shown). This could be due to lack of tissue-specific transcription factors that could be required to express the tissue-specific transcripts. However, according to the HT12 array, the maternally expressed genes are seen to have little to no change in expression. According to Thürmann *et al* (2018) methylation at the *NLRP2* promoter site negatively correlated with expression of the gene as described in our long-term depletion of DNMT1 loss of methylation at the novel *NLRP2* DMR was observed which resulted in the subsequent overexpression. The depletion of DNMT1 causes loss of methylation at the SNURF/SNPRN gDMR and the unexpected repression of the SNRPN. Further to this, the loss of methylation at the SNRPN/SNURF gDMR has been shown to repress the many nearby paternally expressed genes and overexpress the two maternally expressed genes *UBE3A* and *ATP10A* (Court et al., 2014; Horsthemke and Buiting, 2006; Sharp et al., 2010). To further investigate this meta-stable switch observed at the tested imprinted genes it would be interesting to confirm the expression of nearby genes known to be controlled by the respective gDMRs.

Initially, the only DMR at the *NLRP2* loci was described by Woodfine *et al* (2011) with no definition of when the DMR was established, due to the initial investigations in our studies we investigated as it was a possible sDMR. Further to this the trend we seen at the gDMRs and sDMRs led to investigations into the promoter region as a gDMR. To our knowledge, this region has not been previously described as a DMR. Using uniparental ESCs 450k array and RNA-seq datasets this novel DMR appears to inherit methylation paternally resulting in maternal expression of *NLRP2*. The evidence of NLRP2 imprinting is conflicting with one study stating that NLRP2 was a false positive that was picked up in multiple studies (Baran et al.,

2015). However, studies which agree with the imprinting of *NLRP2* describe maternal expression, as the promoter region was also observed to be negatively correlated with expression it would make sense for the promoter region to contain a monoallelic methylation pattern correlating to this maternal expression (Bjornsson *et al.*, 2008; Mahadevan *et al.*, 2017; Thürmann *et al.*, 2018; Pilvar *et al.*, 2019). The novel DMR was also investigated using tracks on UCSC for preferential histone marks, for example, active histone modifications such as H3 and H4 acetylation and H3K4me2/3, an on the inactive allele active histone modifications such as H3K27me2/3 and H3K9me2/3, this novel DMR observed none of these histone modifications which is uncharacteristic for imprinting DMRs (McEwen and Ferguson-Smith, 2010). While not definitive of DMRs CTCF binding and ZFP binding was also investigated as they are found regularly at some imprinted DMRs, CTCF showed no binding at the DMR while some ZFP57 binding was found upstream of the DMR (Singh et al., 2012; Takahashi et al., 2019). This DMR while showing methylation patterns indicative of an imprinting DMR shows no other characteristics of being a DMR.

The *KvDMR* imprinted cluster could be a future region of interest, as according to the 450K array the gDMR known as *KvDMR*, appears to be hypomethylated while the *CDKN1C* DMR appears to be hypermethylated. This is a well-characterised imprinting cluster that is associated with the imprinting disorder Beckwith-Wiedeman syndrome (Smilinich et al., 1999). As mentioned earlier the *Cdkn1c* promoter site in mouse has been classed as an sDMR and has been found to be dependent on the *KvDMR*, though the methylation establishment has not been characterised in human (Bhogal et al., 2004; Woodfine et al., 2011). Another well-characterised imprinting cluster is the *MEG3/DLK1* cluster which is associated with the imprinting disorder Temple syndrome (Ioannides et al., 2014). It can be seen that the *DLK1* and *MEG8* sDMR are both hypermethylated, the nearby gDMR would be the *IG-DMR* but unfortunately, the 450k array does not have probe coverage for this region and would, therefore, need to be tested using a wet lab technique like pyrosequencing. Investigation of these imprinting clusters may give further insight into the mechanism of the “meta-stable” switch between the gDMRs and sDMRs observed in the long-term depletion of DNMT1 in a chromosomally stable human fibroblast cell line.

The Walsh lab recently showed that siRNA and shRNA depletion of DNMT1 results in an overall loss of methylation, with the shRNA showing gains of methylation (O'Neill et al., 2018). This showed that the long-term culture of hTERT-1604s with shRNA targeting DNMT1 causes unexpected gains of methylation, however, these gains were associated with particular chromatin marks. In the current study, there were no consistent chromatin marks that could be associated with the gains seen at the imprinting DMRs when investigated. Globally most imprinting DMRs lose methylation in the long-term depletion of DNMT1 with some regions maintaining or gaining methylation compared to the WT. In comparison, however, in a transient depletion of DNMT1, the imprinting DMRs showed widespread loss of methylation and no gains were observed. This suggests that the initial depletion of DNMT1 caused widespread loss of methylation, however, as the cells continued to be cultured it resulted in gains of methylation at sDMRs. It has been well established that the loss of methylation at gDMRs caused by loss of DNMT1 cannot be recovered in differentiated cells, therefore, the gains of methylation seen at the sDMR regions may be a "coping" mechanism in an attempt to normalise the expression of the imprinted genes as most sDMRs are present at promoter regions such as what was observed with *GRB10* (Chen et al., 2003). A further explanation to the compensation is a similar model suggested by Wang *et al.*, (2019) in which the hypomethylation caused by the depletion of DNMT1 may result in the epigenetic redistribution of nearby histone marks from the sDMR to the gDMR which cannot be restored without germline passage(Thakur et al., 2016). This epigenetic compensation could explain the silencing of SNRPN while the DMR is hypomethylated and the exchange of histone marks from the sDMR may explain why the sDMR retains methylation in the longterm depletion of DNMT1. Redistribution of histone marks is a consistent observation in cultured cells with widespread loss of DNA methylation (Cooper et al., 2014; Ohtani et al., 2018; Walter et al., 2016; Wang et al., 2019). Though this is purely speculation and would require further testing to investigate if this can explain the meta-stable switch.

Not much is currently known about the functionality of many of the sDMRs, The human *GRB10* sDMR, is usually hypomethylated in somatic tissues, except for adult brain in which methylation levels of 35-65% are observed which are levels seen which are typical of a DMR according to Woodfine *et al* (2011) representing that one allele is hypermethylated whilst the other one is hypomethylated. The WT hTERT-1604s *GBR10* sDMR has similar methylation

levels to that observed in the adult brain. Additionally, the potential sDMR described at ATP10a the sDMR has only been observed with imprinted like methylation in the placenta. Finally, the DMR at NLRP2 described by Woodfine *et al* (2011) appears to be hypomethylated in the WT hTERT-1604, it has been previously observed to have methylation levels around 35-65% in several different somatic tissues (Woodfine et al., 2011). Many sDMRs appear to be tissue-specific which may explain the lack of methylation seen at this *NLRP2* DMRs in the WT hTERT-1604. According to the results here, it appears that the long term depletion of methylation at the gDMRs appear to result in the gains of methylation at the sDMRs in a differentiated human cell line, suggesting that the gDMR may play a role in maintaining normal levels of methylation at the sDMRs in human fibroblasts.

4.6 Summary

These results show that the long-term depletion and transient depletion results in different effects on methylation at imprinting DMRs. The long-term depletion of DNMT1 results in a meta-stable switch of methylation between select DMRs defined as gDMR and the sDMR while the siRNA results on the widespread loss of methylation. The meta-stable switch of methylation observed in the long-term depletion of DNMT1 has a functional effect causing aberrant expression of imprinted genes, while previously we have shown that reprogramming can occur with KO and rescue of Dnmt3a2 in mESCs, here we observe to a certain extent reprogramming of imprinted DMRs in a DNMT1 depleted human differentiated cell line. Understanding the mechanism of this meta-stable switch may provide insight into the functionality of sDMRs and further the understanding of the regulation of imprinted genes. Further to this, we investigate a potentially novel DMR that appears to control the monoallelic expression of *NLRP2*.

Chapter 5: UHRF1 and its role in maintaining methylation at imprinted genes and transposable elements

5.1 Introduction

5.1.1 The role3 of UHRF1

The domains and functions of UHRF1 were described during the introduction however the domains will be expanded further in this chapter. Recent research into *UHRF1* has been investigating its potential as a biomarker and a target for cancer therapy even being referred to as a "universal oncogene" (Polepalli et al., 2019). This is not surprising as UHRF1 is essential for the G1/S transition in the cell cycle and is seen to be highly expressed in cells undergoing proliferation (Mousli et al., 2003). Furthering this UHRF1 has also been observed to have important roles in epigenetic inheritance by recruiting DNMT1 and is involved in silencing tumour suppressor genes during cancer progression (Babbio et al., 2012; Sharif et al., 2007). One of the most distinguishing abilities of UHRF1 is that it can link DNA and histone modifications through its SRA and TTD domains respectively (Bronner et al., 2013). Due to the important functions carried out by UHRF1, it is not unexpected that the deletion of *Uhrf1* in mice causes a similar loss of methylation as the deletion of Dnmt1 (Muto et al., 2002; Sharif et al., 2007).

As mentioned previously in this thesis, UHRF1 is made up of 5 domains the UBL, TTD, PHD, SRA and RING domains. The UBL domain is suggested to be involved with the proteasome pathway, the TTD and PHD domains act in conjunction to read the histone H3 code, while the SRA has been shown to elicit a flipping mechanism of the methylated cytosines and finally, the RING domain exhibits E3 ubiquitin ligase activity (Bronner et al., 2013). Until recently there were not many mechanistic studies carried out with disruptive mutations at the domains of *UHRF1* in human. However, a recent study by Kong et al (2019) carried multiple exogenous disruptive point mutations in colorectal cancer (CRC) cell lines depleted of UHRF1. This study revealed that mutations in the TTD domain are unable to undo the hypermethylation observed in CRC, whereas the mutations at the PHD and SRA domains caused the loss of cancer-specific DNA methylation comparable to that of depletion of UHRF1.

The authors suggest that the loss of cancer-specific methylation is due to the disruption of binding between UHRF1 and the histone H3 tail and ultimately interrupts UHRF1's ability to bind to hemimethylated DNA (Cheng et al., 2013; Harrison et al., 2016). Finally, Kong *et al* (2019) also showed that a mutation in the RING domain resulted in a ∼17%–24% loss of methylation in CRC cell lines.

5.1.2 UHRF1 and imprinting

As UHRF1 plays such an important role in the regulation of DNA methylation it comes as no surprise that disruption of *UHRF1* leads to the subsequent loss of imprinting. Sharif et al (2007) showed that the deletion of *Uhrf1* in mESC lead to the loss of methylation at a number known imprinting DMRs. Expanding from this Qi et al (2015) rescued *Uhrf1* null mESC and observed a non-germline restoration of imprints in a small subset of imprints unlike the widespread recovery of imprinted marks observed in *Dnmt3ab* null mESC rescued with *Dnmt3a2*. This suggested the possibility of the *de novo* enzymes playing a more prominent role in the reprogramming of the imprints than DNA maintenance enzymes, as our similar experiment with *Dnmt1* showed no recovery of the imprints (Thakur et al., 2016).

It was later shown that *Uhrf1* null oocytes, when fertilised, would result in loss of methylation at ICRs during preimplantation development, further suggesting that maternal Uhrf1 is essential for ICR maintenance, though showed little to no effect at the ICRs during oocyte development (Maenohara et al., 2017). A more recent study, in contrast, observed that the *Uhrf1* null oocytes do sustain reduced methylation at maternal ICRs as well as in pre-implantation embryos, by showing a loss of methylation at *Peg1* and *Peg3,* with the authors suggesting that the loss of imprinting observed in pre-implantation embryos could be an inheritance from the *Uhrf1* null oocytes (Cao et al., 2019b). However, it should be noted that Maenohara *et al* (2017) examined a larger set of imprints than Cao *et al* (2019) and showed no or insignificant loss of methylation in *Uhrf1* null oocytes, possibly suggesting that *Uhrf1* has no vital role in *de novo* methylation at the imprints and that the loss of maternal Uhrf1 in pre-implantation embryos is what impacts the global reprogramming and subsequent loss of imprinting.

Gene knockout experiments in mice have recently implicated the Krüppel-associated box-containing zinc-finger protein (KRAB-ZFP) ZFP57 in the establishment and maintenance of several imprinted loci (Li et al., 2008) and loss-of-function mutations in the human Zfp57 gene are associated with hypomethylation at multiple imprinted regions in individuals affected by transient neonatal diabetes (Mackay et al., 2008). KRAB recruits KRAB-associated protein 1(KAP1), which acts as a scaffold for various heterochromatin-inducing factors, such as the histone methyltransferase SETDB1 (Schultz et al., 20001, 2002). Therefore, the binding of KRAB-ZFPs to specific chromosomal loci results in nearby transcriptional repression and establishes heterochromatin marks, such as H3K9me3. Correspondingly, KRAB-ZFPs can no longer mediate transcriptional repression when KAP1 is inactivated (Groner et al., 2010). Quenneville et al (2011) associated Uhrf1 with this complex by observing that it coimmunoprecipitates with KAP1 in mESCs. Further to this differential chromatin marks has been observed at the active and inactive alleles at imprinted genes, in particular, H3K9 is widely associated with the silent allele of imprinted genes (Chang and Bartolomei, 2020). The UHRF1's binding to H3K9 and it's association with the ZFP57/KAP1 complex it is not a surprise that loss of UHRF1 results in the loss of imprinting.

5.1.3 UHRF1 and retrotransposons

Similar to ICRs, some transposable elements are also protected from demethylation through reprogramming during implantation (Reik et al., 2001). Interestingly a small number of imprinted genes resemble retrotransposons or have even been passively retrotransposed themselves (Cowley and Oakey, 2010). Further to this ERV's and inactive imprinted alleles contain both the repressive mark H3K9me3 (Matsui et al., 2010). While seemingly widely different in their function and regulation it has been predicted by Barlow (1993) that imprinted genes may have arisen from a host defence mechanism to silence retrotransposons. While this was predicted 27 years ago, many of the predictions are still valid to this day, such as various maternal and paternal imprinting factors, imprinting boxes with the presence of a key sequence element containing 'foreign invading DNA' and a host system to neutralise the key sequence elements resulting in identical machinery used for repression at these elements (Ondičová et al., 2020). The KRAB-ZFP complex discussed earlier is one such machinery that is linked with the repression of transposable elements and imprinted gDMRs,

this machinery works in a sequence-specific manner which is attributed to the unique combinations of zinc-fingers (Monteagudo-Sánchez et al., 2020). In particular ZFP57 one of the most studied members of this complex has been shown to target IAP, LTR and imprinted gDMRs (Li et al., 2008; Shi et al., 2019). The mapping of 200 KRAB-ZFP DNA binding proteins target shows was predominantly found at transposal elements including ERV's (Imbeault et al., 2017). Similarly to the mechanism observed at the inactivated alleles at imprinted regions, the retrotransposable elements once bound by KRAB-ZFP binding resulting in the recruitment of KAP1 and other proteins involved in the complex such as SET domain bifurcated 1 (SETDB1) and UHRF1 ultimately inducing local acquisition of H3K9me3 and DNA methylation (Quenneville et al., 2011; Shi et al., 2019). While studies investigating UHRF1 in regards to the maintenance of imprints appear to agree, the role of UHRF1 in maintaining transposable elements is clouded with conflicting literature.

Mutations in UHRF1 were initially characterised as phenocopying loss of DNMT1 in mouse and resulted in widespread hypomethylation of the genome and dysregulation of imprinted genes, as well as ERV such as IAP (Bostick et al., 2007; Sharif et al., 2007). Mutations in the TTD-PHD region that affect H3K9me3 binding by UHRF1 have been shown in human to decrease DNA methylation at ribosomal DNA repeats in HeLa cells (Rothbart et al., 2012). A further study also observed an upregulation of endogenous retroviruses as well as up-regulated expression of viral defence genes in a colorectal cancer cell line with a long-term depletion of UHRF1, a response which was greater than DNMTi treatment in the same cell line (Cai et al., 2017). The same group went on to show that point mutations in the PHD and SRA domain lead to similar demethylation and derepression of the endogenous retroviruses (Kong et al., 2019). In mouse, mutations in the same region gave only a 10% decrease in DNA methylation, which was genome-wide and not just restricted to repeats (Zhao et al., 2016). Controversially, Sharif et al (2016), contradicting a paper published previously from their lab, stated in their study from 2016 that repression of transposable elements requires both methylation-dependent and methylation-independent pathways, suggesting the activation seen in *Dnmt* null mice was due to prolonged binding of UHRF1 preventing the establishment of H3K9me3 by the H3K9 methyltransferase SETDB1, rather than the loss of DNA methylation. In contrast, mutations in the zebrafish homologue were reported to result in ERV derepression in the developing embryo and activation of the innate immune system as for

DNMTi in human, but through double-stranded DNA rather than dsRNA signalling (Chernyavskaya et al., 2017). There is, therefore, a lack of clarity regarding the role of UHRF1, what the cellular response to the loss of this important epigenetic regulator would be, what genes would be most affected and what the dependence, if any, of DNA methylation on the TTD-PHD domain would be during development.

5.1.4 Stable knockdown of UHRF1 in hTERT-1604s global loss of methylation and de-repression of ERVs

The preliminary study carried out by the Walsh lab was carried out in a normosomic differentiated fibroblast cell line hTERT-1604 (described earlier in this thesis) which had been depleted of UHRF1 (UH4). The UH4s were generated in the same fashion as shown in the earlier research chapter for the D8, D10 and D16 cell lines (Figure 5.1A). The Walsh lab has previously confirmed the depletion of UHRF1 in this model (Figure 5.1B and C) and subsequent global loss of methylation similar to that of the depletion of DNMT1 (Figure 5.1D). In this investigation, Irwin et al (In Prep) implicated H3K9me3 in ERV de-repression in the absence of DNA methylation and showed that the PHD domain of UHRF1 plays an important role in the repression of ERVs in a human model (Figure 5.1E-G). The implication of the H3K9me3 binding was evident as cell lines rescued with the point-mutated UHRF1 (PHD1, PHD4, PHD10, TTD9) failed to repress the ERV's but when the same UH4 cells were rescued with intact protein (WT10, WT18) repression occurred (Figure 5.1G). However, interestingly in this study by Irwin et al (In Prep), WT10 showed no recovery of methylation at the ERV elements (data not shown).

Figure 5.1

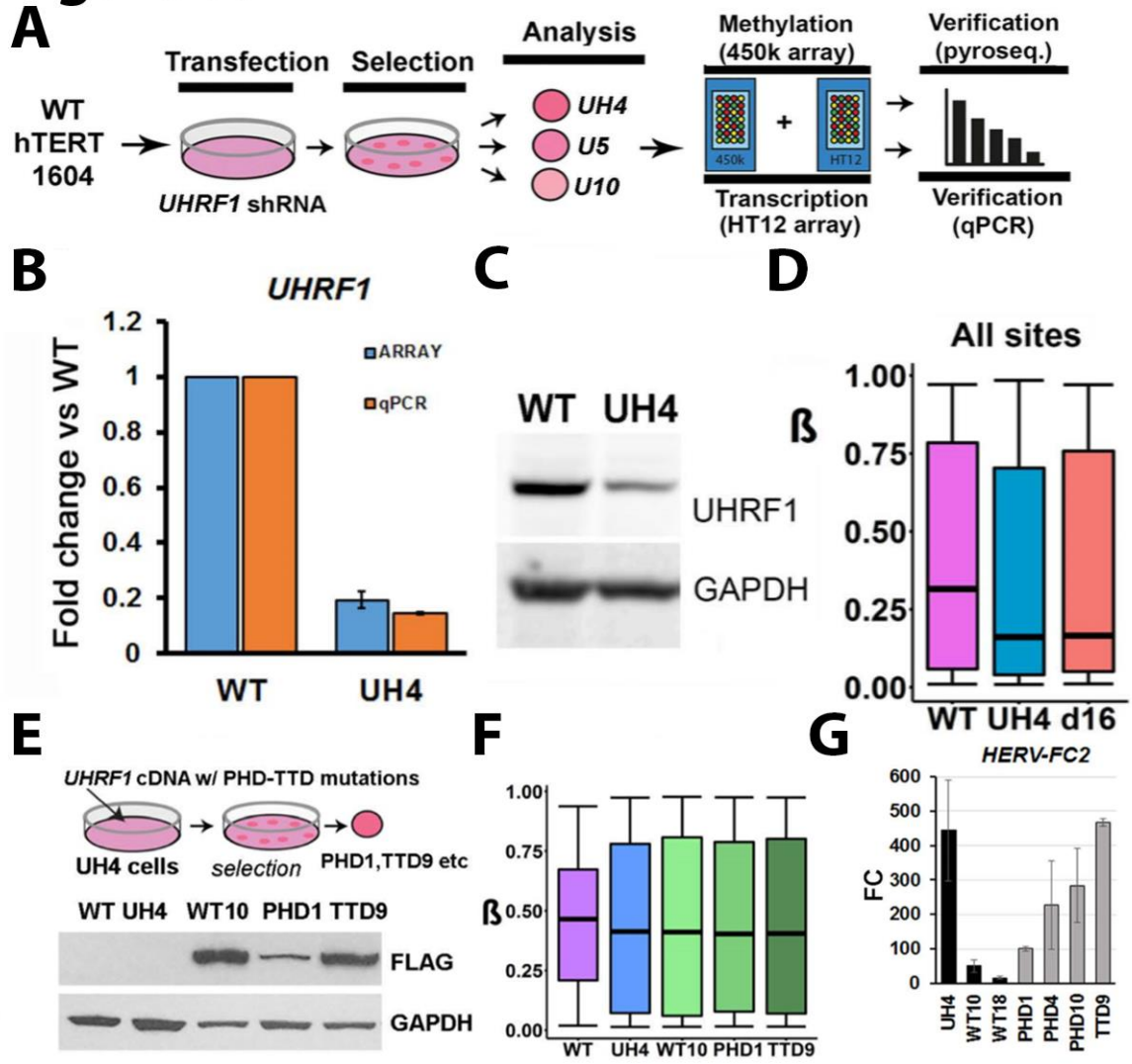


Figure 5.1: Stable UHRF1 knockdown depletion results in a global loss of methylation and rerepression of ERVs in hTERT-1604s. (A) The experimental approach is taken to generate and analyse the UHRF1-depleted hTERT-1604s. Transfection was carried out using an shRNA targeting UHRF1 and then cells grown in selective media until colonies appeared. Three separate clonally-derived cell lines were expanded (UH4, U5 and U10) and further analysed with methylation and transcription arrays and confirmed with wet-lab techniques. (B) Confirmation that UHRF1 expression is depleted in the UH4 cell line using the HT12 array and qPCR and **(C)** loss of protein using western blot. **(D)** UH4 and d16 show similar changes of methylation across all probes on the Infinium human 450k array: a βvalue of 1 equates to 100% methylation. Median values are indicated by the line and whiskers represent interquartile range. **(E)** The experimental approach is taken to generate rescue cell lines with UHRF1 proteins including proteins containing point mutations previously shown to affect binding (WT10, WT18, PHD (D334, E335) and TTD (Y188)) and colonies expanded. **(F)** All rescue cell lines fail to recover methylation and shows similar changes of methylation across all probes on the Infinium human 450k array. **(G)** qPCR shows derepression in UH4s in comparison to WT, while the rescue cell lines WT10 and WT18 repress the reactivation similar to WT levels, however, when UH4 is rescued with the point mutated proteins PHD1, PHD4, PHD10 and TTD9 the extent of repression is not re-established to the same level as what is seen in the WT protein. Data generated by Drs. Avinash Thakur Rachelle Irwin and Catherine Scullion

It is clear that the loss of UHRF1 results in the rerepression of retrotransposable elements and loss of imprinted expression. We hypothesize that a mutation within the PHD domain that is essential for the binding of UHRF1 to H3K9 will result in the rerepression of retrotransposable elements and loss of genomic imprinted expression without effecting the Uhrf1 expression.

5.2 Aims

The overall aim of this chapter is to provide more insight into the importance of the PHD domain in Uhrf1. Previously the Walsh lab had shown de novo methylation activity in hTERT-1604 cells was sufficient to restore methylation to WT at some genes (O'Neill et al., 2018). We considered that these adult cells may instead lack other factors required for ERV DNA methylation which are only found earlier in development. Therefore, to examine the dependence of de novo methylation and ERV repression on an intact H3K9me3 binding domain in UHRF1, we generated mouse embryos containing mutations in the PHD domain matching those used in the hTERT-1604 UH4 cell line. We will further this investigation to include the genomic imprinted regions due to the overlapping similarities between transposable elements and the genomic imprinted regions using both the PHD mouse mutant embryo and hTERT-1604 models.

- Firstly, we will generate a mouse model with a *Uhrf1* loss of function PHD mutation to investigate the dependence of an intact H3K9me3 binding domain in UHRF1 for de novo methylation and retrotransposal repression during embryonic development.
- Next, we will further explore the requirement of the PHD domain for the maintenance of genomic imprinted elements during mouse embryonic development.
- Finally, we will investigate the consequence of the long-term depletion of UHRF1 has on the retrotransposal and genomic imprinted elements in a human chromosomally stable cell model

5.3 Materials and Methods

Methylation analysis was carried out with the Infinium 450k beadchip array and pyrosequencing while expression analysis was carried out with qPCR as described in chapter 2. Further investigations were carried out for methylation and expression using pyrosequencing, clonal analysis and *in situ* hybridisation.

5.3.1 Pharmacological inhibition of DNMT1

Cell culture was carried out as described in Chapter 2. For drug treatment, 5-Aza-2'-deoxycytidine (5'Aza-dC; Sigma-Aldrich, Dorset, UK) was dissolved in DMSO and added to culture media; negative controls contained just DMSO. Cells were washed twice with PBS before the addition of 10mls fresh complete media supplemented with 5-Aza-2'-deoxycytidine (5'Aza-dC; Sigma-Aldrich, Dorset, UK) at a final concentration of 1µM. Cells were then treated with 5-AZA-CdR at a final concentration of 300nM at 0 hours. After 24 hours the cells were washed with PBS and fresh media without drug was replaced. Cells were grown for up to 36 days with the media changed every two days. Cells were harvested at 72 and 96 hours.

5.3.2 CRISPR-Cas9 generation of Uhrf1 PHD D334/E335AA mutant mice

Superovulation was induced in 5-week-old B6D2F1 female mice through intraperitoneal injection of pregnant mare's serum gonadotropin (5 IU). 48 hours later the mice were injected with chorionic gonadotropin (p5 IU) then mated with male B6D2F1 mice overnight. After which zygotes at Embryo day (E) 0.5 were collected from the oviducts of female mice. To remove the cumulus cells zygotes were incubated in 1% hyaluronidase/M2 medium followed by washing with fresh M2 medium and recovery for 6h at 37 °C in a 5% CO_2 incubator. Microinjection containing a mixture of SpCas9 mRNA (100 ng/µl), *Uhrf1*-crRNA (50ng/µl) and 112-bp single-stranded oligodeoxynucleotide (ssODN, 10ng/µl), which was flanked by homologous arms corresponding to exon 7 of *Uhrf1 was carried out using* a FemtoJet microinjector, set to Pc = 10-15 hPa and Pi = 40-50 hPa. Zygotes that were successfully microinjected were then incubated for 72h in KSOM at 37 °C in a 5% CO_2 incubator until they reached the blastocyst stage when they were then transferred into the

uteri of pseudopregnant female *B6D2F1* mice. To confirm the CRISPR/Cas9 Uhrf1 mutation PCR using specific primers (Table 1) was carried out to amplify the genomic regions flanking the gRNA target. The PCR amplicons were ligated into the pClone007 vector and sequenced.

5.3.3 Cloning

The primers IAP-F 5'-ttg ata gtt gtg ttt taa gtg gta aat aaa-3' and IAP-R 5'-caa aaa aaa cac cac aaa cca aaa t-3' were used to generate a PCR product from bisulphite converted DNA following the pyro PCR protocol in chapter 2. The PCR product was cloned into the cloning vector pClone007 using the pClone007 Simple Vector kit according to the manufacturer's instructions (Beijing Tsingke, Beijing, China). 1ul of vector plasmid was then used to chemically transform competent E.coli cells, by heat shocking the E.coli cells at 42^{o}C for 45 seconds. E. Coli cells were then incubated in Lennox broth (Sigma-Aldrich, St. Louis, Missouri, United States) for 1 hour then plated out on Luria-Bertani (LB) agar (Shanghai shisheng Sibas advanced technology, Shanghai, Chin) agar plates containing 50mg/ml ampicillin (Shanghai shisheng Sibas advanced technology, Shanghai, China) for selection of transformants. Plates were incubated overnight at 37^{o}C. Agar plates containing colonies were then sent to Beijing Tsingke (Beijing, China) for plasmids to be extracted and sequenced.

5.3.4 Bisulphite Sequencing

The DNA sequence for the area of interest was obtained, bisulfite converted and aligned with sequencing results using the EMBL-EBI Clustal Omega Multiple Sequence Alignment tool; with sequences inputted as FASTA format. The methylation status at each CpG site was confirmed by the retention of the cytosine of CG sites while the gain of thymine indicated the loss of methylation at the CG site when compared to the reference sequence.

5.3.5 In situ IAP probe preparation

To label the IAP in mouse embryos an RNA probe was generated through in vitro transcription-based on previously used primer sequences (Walsh, Chaillet & Bestor, Nat Gen 1998). An IAP probe template was generated using PCR, the PCR was set up using 25ul 2×Gflex PCR Buffer (Mg2 and dNTPs), IAP in situ primers -A45:

'GGGCCCAATACGACTCACTATAGGGAGGCCAGAATCTTCTACGGC' which contains a T7 promoter to transcribe the RNA probe containing DIG-labelled nucleotide *in situ* and B-17 'CTGTAAGCTTAAGGCCC' at a final concentration of 0.3µM, 500ng of mouse embryo DNA template and 1 µl of Tks Gflex DNA Polymerase (1.25 units/µl) all reagents were sourced from Takara Bio (Shiga, Japan). PCR was carried out using the following conditions: 1 min at 94 °C followed by 35 cycles of 98°C for 10s, 60°C for 15s, 68°C for 35s and a final elongation step of 68°C for 3min. The IAP template was then purified using the QIAquick Gel Extraction Kit (Qiagen, Crawley, UK) according to the manufacturer's instructions. A DIG-labelled, single-stranded RNA probe was then generated using the DIG RNA Labelling Kit (SP6/T7) (Roche, West Sussex, UK) by following the manufacturer's instructions.

5.3.6 Whole-mount in situ hybridization

Embryos were prepared by fixing in 4% paraformaldehyde in phosphate buffer saline (PBS) at 4°C overnight. After the overnight fixation embryos underwent washing with PBS/Tween (PBT) (1 × PBS plus 0.1% Tween) for 5 minutes with gentle rocking at 4°C. The embryos were subjected to a dehydration series using a methanol/PBT series of 25%, 50%, 75% and 100%, each stage lasted 5 minutes and was carried out at 4°C with gentle rocking. A final 100% ethanol step was carried out with the same conditions, once this was done the embryos were stored up to 1 month in the -20°C. The embryos were removed from storage and rehydrated to begin the hybridization, firstly rehydration was carried out with a series of methanol/PBT of 75%, 50% and 25% with the same conditions as the dehydration series. A bleaching step was conducted out with 6% hydrogen peroxide in PBT at 4°C for 1 hour. The embryos underwent to washing steps with PBT for 5 minutes with gentle rocking at room temperature. The embryos underwent proteinase K treatment at 10g/ml in PBT for 6 minutes at 4°C. Then followed up with a wash of 2mg/ml glycine in PBT for 5 minutes followed by another two washes with PBT for 5 minutes each. Afterwards, the embryos were washed in increasing concentrations of hybridization solution (50% formamide, 5 × SSC, pH 4.5 (citric acid to pH), 1% SDS, 100 µg/ml yeast tRNA, 50 µg/ml heparin) and PBT mixture with the ratios of 1:3, 1:1, 3:1, the embryos were washed for 5 minutes in each mixture, this was followed by a 10-minute wash in the hybridization solution. The hybridization solution was replaced with fresh hybridization solution and incubated at 65°C for 6 hours and then the hybridization solution

was replaced with fresh solution for a final time but also contained 200ng/ml of the IAP probe and incubated at 65°C overnight. Proceeding to the overnight incubation the embryos were washed in a 1:1 mixture of hybridization solution and FA/SCC/Tween (FST) (50% formamide 5 × SSC, 0.1% Tween and DEPC water; at pH 4.5) for 30 minutes then a 30-minute wash in FST at 65°C. The embryos were then allowed to cool down for 10 minutes at room temperature in a 1:1 mixture of FST and Tris-buffered/Saline/Tween (TBST) (NaCL (FW 58.44), KCL (FW 74.55), 1M Tris-HCl, Water; at pH 7.5). The embryos were blocked using a blocking reagent (Roche, West Sussex, UK) for 1 hour at room temperature, the blocking reagent is then replaced with fresh reagent with the addition Anti-Digoxigenin-AP Fab fragments (Roche, West Sussex, UK) (1:2500) and incubated at 4°C overnight. The next day the embryos were washed 3 times for 5 minutes in TBST, followed by 5 more times for 1 hour in TBST and then a final wash of TBST overnight at 4°C overnight with gentle rocking. After the overnight incubation in TBST the embryos were washed 3 times in NaCl/Tris–HCl/MgCl/Tween (NTMT) (100 mM Tris-HCl, 100 mM NaCl, 50 mM MgCl2, 0.1% Tween 2 mM (0.5 mg/ml) Levamisole; 9.5pH) for 20 mins at room temperature. The NTMT was then removed and an NBT/BCIP stock solution (Roche, West Sussex, UK) made up with NTMT was added and left at room temperature for 40 minutes to allow for the colour to develop and was stopped by washing with PBS and then fixed using 4% paraformaldehyde/ 0.1% glutaraldehyde, before two final washes of PBS were carried out.

Table 5.1 RT-qPCR primers used in this chapter

Gene	Primer	Oligo sequence (5’-3’)
Human		
GRB10	FWD	GCACGAAGACAACCAGGTG
	REV	GAACGATCACTGCCTTACCC
SNRPN	FWD	CACCAAGAGGTGGTTAAAGC

	REV	GATTGCTGTTCCACCAAATCC
NLRP2	FWD	CGGCCGAGAGAGAAGCCTTATTAG
	REV	CAACACCTGGCCCTACTCGC
UHRF1	FWD	TGAGGACATGTGGGATGAGA
	REV	GTCCCTGGAGTTCATCTGGA
IFI27	FWD	TCTGTCCACCCTCTGCTTCT
	REV	GGCATGGTTCTCTTCTCTGC
BST2	FWD	AGGGAGGTGTCATCGTCAAC
	REV	CTCTTGTCTCCCACCCTGAG
HERV-FC2	FWD	TTTCCCACCGCTGGTAATAG
	REV	AGGCTAAGGATTCGGCTGAG
STAT1	FWD	CCCACTCTGATCAACTTTTGC
	REV	GGCCTGTTGAAGATGCTTGT
Mouse		
H19	FWD	CGATTGCACTGGTTTGGA
	REV	CTCAGACGGAGATGGACGA

Igf2	FWD	GGATCCCAGAACCCGAGAAGA
	REV	GGGCGGCTATTGTTGTTCTCA
Grb10	FWD	GTAAGCGGAGCACACGGATGAA
	REV	TACGAACGCCTTTGGATTACTCTG
mUhrf1 PHD mutant	FWD	TGACTCTCAGCTCAACAACTGTCG
	REV	CTCTCACTGAGCCTACAGCCAAG

Table 5.2 Bisulphite primers used in this chapter

Gene	Primer	Oligo sequence (5'-3')
Human		
GRB10 gDMR	FWD	GGGAGAAAGAGGTTTTTA
	REV (Biotin)	AAATCTAAACATCC
	SEQ	ATTTAAAAAATAAATAAATCTAAACATCC
SNRPN gDMR	Qiagen Assay	Hs_SNURF/SNRPN_01_PM PyroMark CpG assay (PM00168252)
NLRP2 gDMR	FWD	GGGATTGGTTTGAATTGTAG

	REV (Biotin)	CCTTAAATAACATCACCCTATTCAACA
	SEQ	ATTGGTTTGAATTGTAGGA
Mouse		
Grb10	FWD	GGGGTTAGTGGATAGTTT
	REV (Biotin)	CTAAACTCCAAAACCCTTTTTCTA
	SEQ	GTTTTAGAATTAGGTAGTATTTAG
Snrpn	FWD	GGTTAGAGGGATAGAGATTTTTGTATTG
	REV (Biotin)	TCCACAAACCCAACTAACCT
	SEQ	TATGTGTAGTTATTGTTTGGGA
H19	FWD	TTTTTTGTTGAATTTGGGGTATTTAAAG
	REV (Biotin)	CACATTTCTTAAATAACTCCTTCAATCTT
	SEQ	ATTTGGGGTATTTAAAGT
KvDMR	FWD	AGGAGGGGGAGGTTATGA
	REV (Biotin)	ACTTTTCTATTCAACTTAATTCCCAAC

	SEQ	GGGGGAGGTTATGAT
Mest	Qiagen Assay	Mm_Mest_01_PM PyroMark CpG assay (PM00384230)

5.4 Contributions

For this chapter, the PHD mutant mouse was generated by Dr. Rachelle Irwin and Meiling Sun. I helped harvest embryos and carry out in situ. Cell lines were generated by Drs. Avinash Thakur and Rachelle Irwin. I revived lines from cryostorage, grew the cells and isolated nucleic acid for analysis. Primer sequences and PCR conditions for some human retrotransposable elements and immune genes were generated by Dr. Catherine Scullion. All genomic imprinted investigations were carried out by myself as well as the clonal analysis for the *IAP* elements in mouse studies and the HT29 5-AZA treatment and investigation.

5.5 Results

5.5.1 Generation of a Uhrf1 PHD mutant mouse line

Generation of a *Uhrf1* PHD mutant mice required three rounds of injections due to low survival rate, suggesting that the PHD mutation had a highly lethal effect in development as described in Figure 5.2A. The morphology of the embryos at e8.75 and E9.5 looks relatively normal for the WT and HET mutant mice, but there are clear morphological abnormalities at E8.75 which is more extreme by e9.5 (Figure 5.2B). The mutation was confirmed using Sanger sequencing (Appendix Figure A1) and was observed with hypomethylation at the *IAP 5'UTR* at 2 individual E8.75 and E9.5 HOM embryos against a WT and HET embryo respectively (Figure 5.2B). Hypomethylation was then further observed at the *IAP LTR* at all 6 CG sites examined using pyrosequencing (Figure 5.2C). Methylation at the repetitive elements *LINE-1* and *IAP* 5'UTR show highly significant loss of methylation (Figure 5.2D). Further to this, the *Uhrf1* PHD mutation leads to derepression at the *IAPez-GAG* and *musD* elements, (Figure 5.2E). A more in-depth transcriptional analysis confirmed derepression across several different retrotransposable elements (Figure 5.2F). While the Uhrf1 results in clear phenotypical defects during embryonic development and rerepression of the transposable elements, it is seen that the UHRF1 expression of the protein remains close to WT levels (Figure 5.2F). The Sanger sequence and sgRNA guide information can be seen in Appendix Figure A1. While generating the mice PHD mutants, the Walsh lab also attempted to generate 2 human HEK293T cell lines with either a *UHRF1* KO or a PHD mutation using CRISPR-Cas9 (Appendix Figure A1). However, this proved to be lethal and resulted in no colonies growing from a single cell for either edit.

Figure 5.2

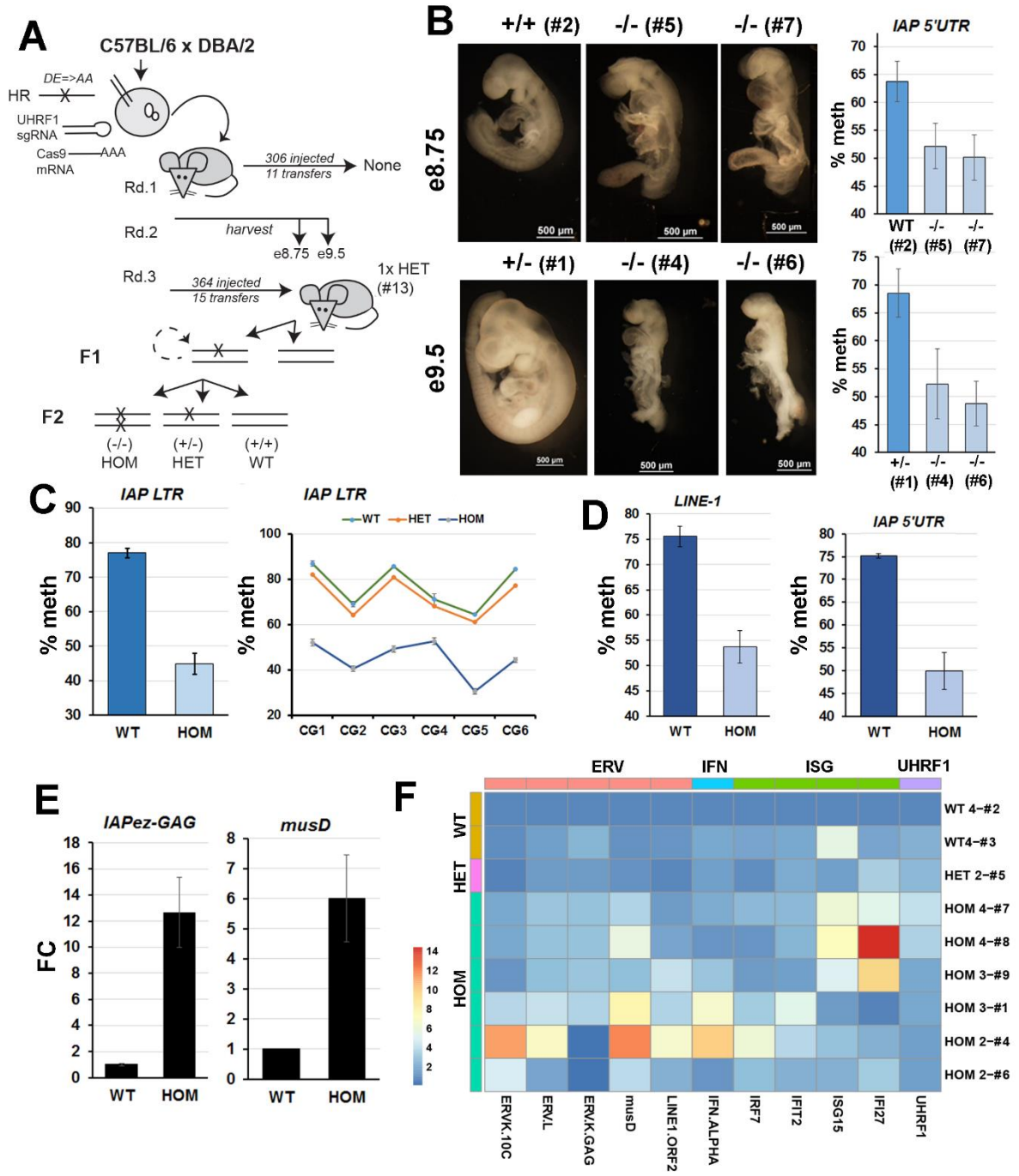


Figure 5.2: Uhrf1 PHD mutation results in loss of methylation and derepression at several transposable elements during embryonic development. Mutation of the PHD domain of Uhrf1 in mouse embryos was carried out using CRISPR. **(A)** The Uhrf1 sgRNA were injected into C57BL/6 and DBA/2 embryos which were then implanted into a female surrogate. The first round resulted in embryo lethality and no offspring, the second round of injections were harvested at e8.75 and e9.5 to be used for investigations and finally the third round of injections resulted in a heterozygotic *Uhrf1* PHD mutant. This HET mouse was then used as a founder for the *Uhrf1* PHD mutation mouse line. **(B)** The mouse embryos for the WT and HET show normal morphology at e8.75 and e9.5 in which they have started turning, meanwhile, at the same age, the homozygous mutants appear smaller and less advanced than their WT and HET counterparts at the same ages. Overall levels of methylation were significantly lower at the *IAP* 5'UTR in both E8.75 and E9.5 Homozygotic mutant embryos. **(C)** Methylation at the *IAP LTR* is

hypomethylated across all 6 CG sites investigated using pyrosequencing. **(D)** Significant loss of methylation was also observed at both the *Line-1* and *IAP* transposable elements. **(E)** *IAPez-GAG* and *musD* both show derepression in homozygotic *Uhrf1* PHD mutants. **(F)** Further, in-depth transcription analysis of retrotransposons shows that a selection of different elements is derepressed in several different homozygotic *Uhrf1* PHD mutants, with blue representing no change of expression and red representing large upregulation of element.

5.5.2 Mouse Uhrf1 PHD mutant and IAP derepression

To further investigate the derepression of the retrotransposable elements alternative methods were utilised. This included clonal analysis for a number of reasons: a) it could confirm methylation loss independently; b) if there was a mixed population of methylated and unmethylated transposons, this could be distinguished using this method; c) sites can be directly compared between clonal analysis and pyrosequencing, which is not always the case and d)since there are many *IAP* elements with different sequences due to mutations, we could get an impression of the extent to which methylation changes were linked to particular types.

Figure A1 shows that all the *IAP* elements investigated using clonal analysis showed the expected heavily methylated CG sites in the WT mouse embryos E8.75 (Walsh et al., 1998). The *IAP* elements in the WT e8.75 mouse embryos contained 97% methylated CpG sites compared to the PHD mutant in which only 35.71% CpGs were methylated on average across all *IAP* elements investigated (Figure 5.3A-B). This confirms the loss of methylation from the pyrosequencing analysis with similar overall losses of methylation. To further gauge the effect on which the *Uhrf1* PHD mutation has on IAP during embryonic development, an *in-situ* analysis was carried out for WT and HOM mutants at e8.75 and e9.5 using a probe specific for a recently active or young IAP. In Figure 5.3C it is clear that the signal is stronger in both e8.57 and e9.5 HOM mutants when compared to the respective WT which is clearer in the latter.

Figure 5.3

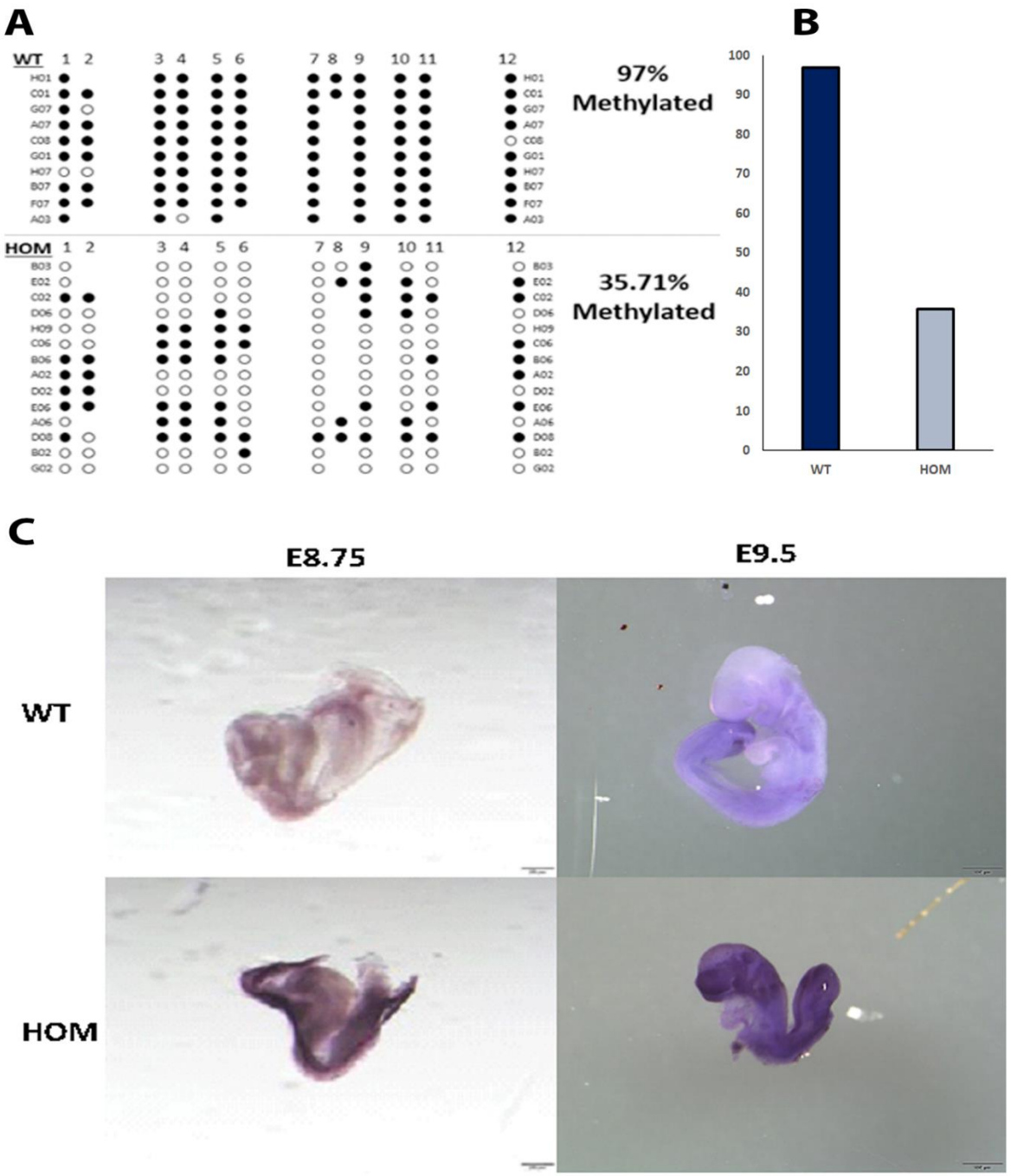


Figure 5.3: Mouse embryos with homozygotic mutations in the PHD domain of Uhrf1 present with loss of methylation at IAP repetitive elements. *IAP* are repetitive elements found in mice, using bisulphite sequencing it is possible to measure the levels of methylation across the different repeats found throughout the genome. **(A)** Clonal analysis showing methylation across the *IAP* sequences, Black circles represent methylated CG sites while white circles represent unmethylated CG sites that have been converted to TG during bisulphite conversion. **(B)** Visual representation of the methylation from the clonal analysis. **(C)** Using a probe targeting IAP an *in-situ* hybridisation was carried out on mouse embryos at ages E8.75 and E9.5 for both wildtype (WT) and homozygotic PHD mutations *Uhrf1* (HOM). There was a clear *IAP* upregulation in the HOM embryos at both ages as shown by the darker colour. However, it is much more prominent in the E9.5 HOM suggesting the derepression is more severe at this later stage of embryonic development in the mutant embryos.

5.5.3 UHRF1 is important for the maintenance of imprinting

Uhrf1 has been previously observed to play a role in the regulation of imprints during embryonic development, here we investigate the potential involvement of the PHD domain of *Uhrf1* in maintaining the imprints during embryonic development. To investigate this, embryos were harvested at e8.75 and used to extract DNA and RNA. As expected, pyrosequencing technology indicated normal methylation (normally 45-55% in the mouse embryo (Thakur et al., 2016)) at the imprinted gDMRs in WT embryos. The PHD mutation, however, resulted in significant loss of methylation across the 4 imprinted DMRs investigated, with the *H19* DMR just missing out on significance with a p-value of 0.0563 (Figure 5.4A). The loss of methylation at the *H19* and *Grb10* ICRs subsequently resulted in the loss of imprinted expression of *H19* and *Igf2*, which are normally controlled by the *H19* gDMR and Grb10 which is normally controlled by the *Grb10* gDMR respectively (Figure 5.4B-D). However, only *Grb10* showed significant changes in expression.

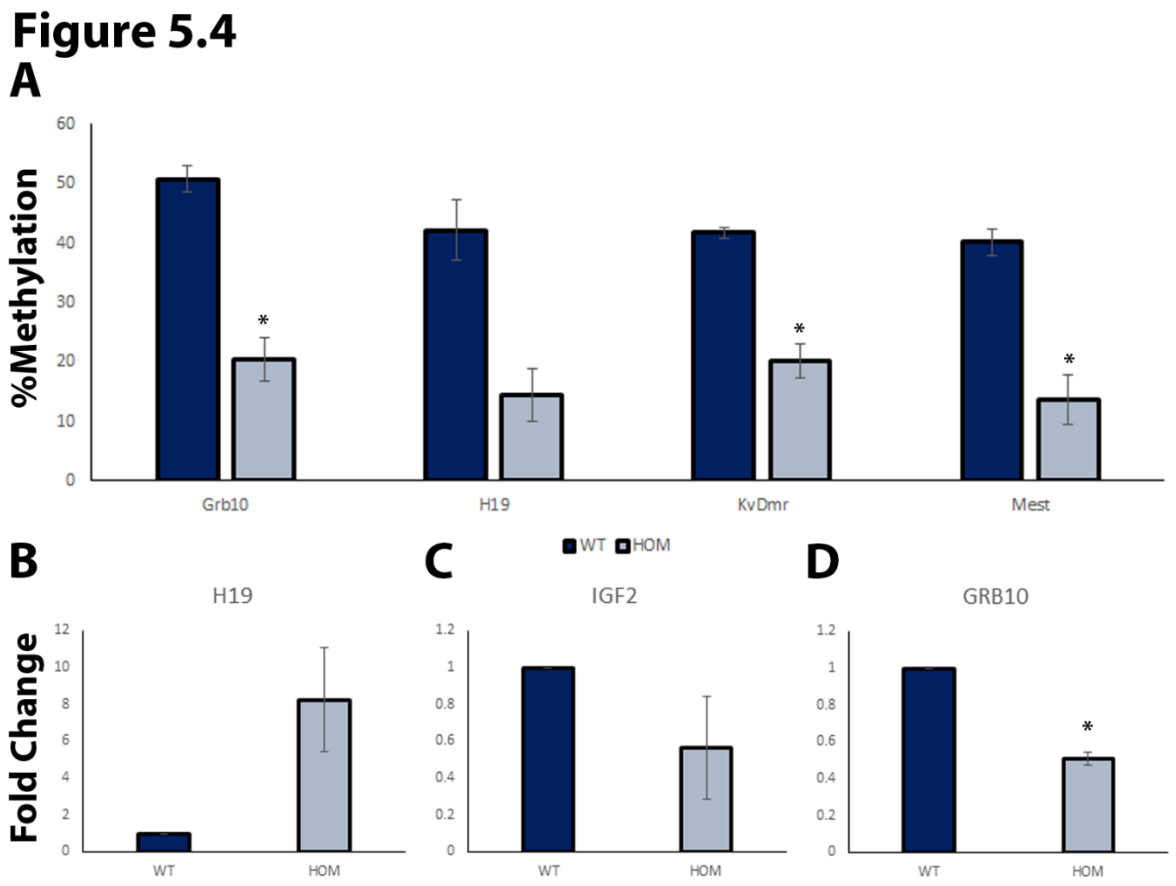


Figure 5.4: Mouse embryos with homozygotic mutations in the PHD domain of Uhrf1 present with loss of imprinting at E8.75. (A) Pyrosequencing confirms the loss of methylation at the imprinted DMRs in mouse embryos E8.75 with a homozygotic mutation at the PHD domain of *Uhrf1* (HOM) when compared to a wild type (WT). This was then seen to have a functional effect with E8.75 HOM embryos which appear to lose imprinted expression at **(B)** *H19*, **(C)** *Igf2* and **(D)** *Grb10*. Due to lack of embryonic material experiments were carried out in duplicate with biological replicates. Error bars indicate standard error of the mean. Significance carried out with a two-tailed T-test with the following key, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

5.5.4 5-aza-2'-deoxycytidine activates 'viral mimicry' response

Initial investigations by the Walsh Lab noted that the UH4 cell line had upregulated expression of genes involved in 'viral mimicry' such as dsRNA sensors, transcription factors, anti-viral response, cell death and T-Cell signal (Figure 5.5A). Roulois et al (2015) have previously shown that the treatment of cells with low-dose 5-aza-2'-deoxycytidine (AZA) results in the activation of 'viral mimicry' in human colorectal cancer cells by inducing the expression of type III interferons (IFNs) and subsequent upregulation of interferon-stimulated genes (ISGs). We then confirmed that low dose treatment of AZA does result in the activation for the viral

defence genes in *STAT1* (Transcription factor), *BST2* (Anti-viral response) and *IFI27* (Cell death effects) (Figure 5.5B-D).

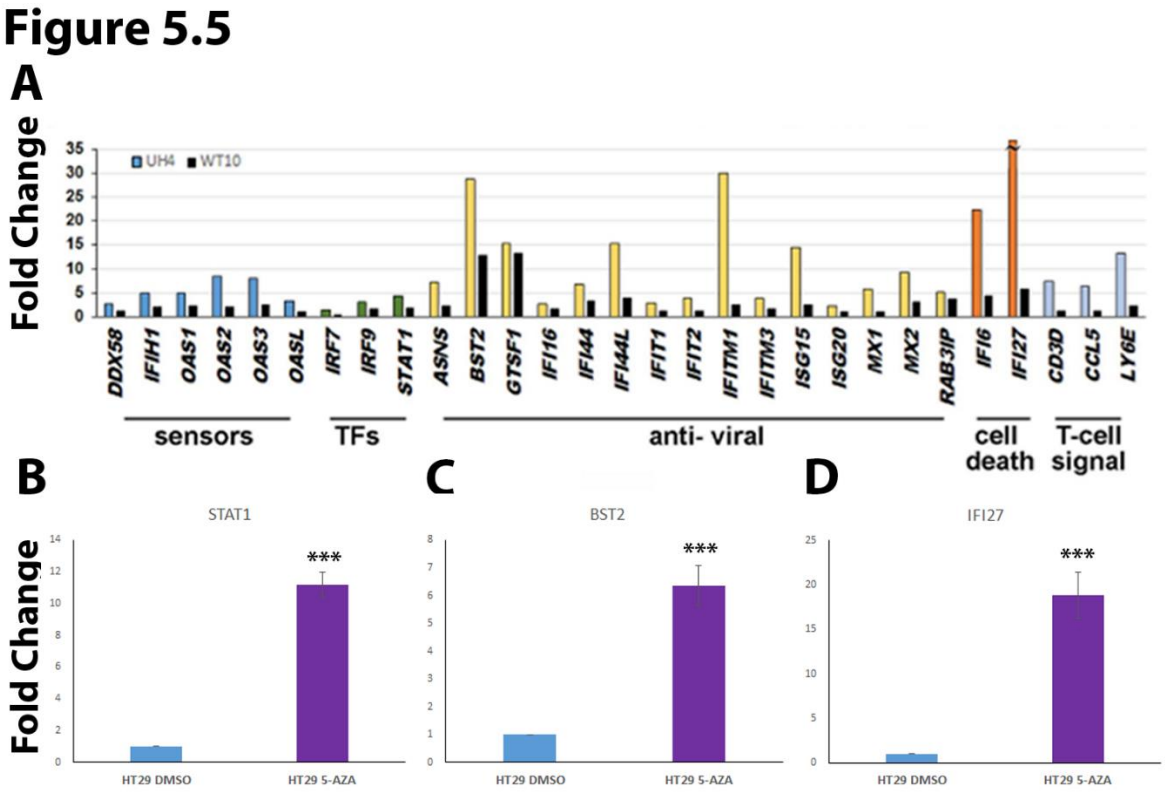


Figure 5.5: Stable UHRF1 knockdown depletion results in upregulation of genes involved in viral mimicry like what is observed in low dose treatment of AZA in human colorectal cancer cells. (A) Average fold change from HT12 transcriptional array show expression of the viral defence genes vs WT cells Data generated by Dr. Catherine Scullion. RT-qPCR expression analysis of HT29 human colorectal cancer cells treated as in 5-aza-2'-deoxycytidine (AZA) for **(B)** *STAT1*, **(C)** *BST2* and **(D)** *IFI27*. qPCR was carried out in triplicate with technical replicates. Error bars indicate standard error of the mean. Significance carried out with a two-tailed T-test with the following key, *p < 0.05; **p < 0.01; ***p < 0.001. TFs; transcription factors, FC; fold change, DMSO; Dimethyl sulfoxide.

5.5.5 Investigating the role of UHRF1 in imprinted maintenance in human

By analysing the UHRF1-depleted cell line UH4 using the 450k Infinium beadchip array, it was clear that methylation was lost at a subset of imprinting gDMRs (Figure 5.6A). The DMRs *GRB10* and *SNURF* showed the largest losses of methylation followed by the *KvDMR, MKRN3_MIR4508* and *ZNF597_NAA60* DMRs. The remaining DMRs analysed only observed small changes of methylation around +/-5%. Interestingly, rescuing the depleted hTERT-1604s with a plasmid containing *UHRF1* (WT10) did not result in a recovery of methylation at the imprinted gDMRs (Figure 5.6A). Further methylation analysis of the *SNURF* and *GRB10* DMRs

with pyrosequencing demonstrated a significant loss of methylation at both gDMRs in the UH4s and WT10 in comparison to the WT hTERT-1604s (Figure 5.6B). Loss of methylation observed at the UH4 imprinting DMRs *GRB10* and *SNURF* resulted in significant loss of expression for both GRB10 and SNRPN in comparison to the WT hTERT-1604s. The imprints additionally failed to recover expression in the WT10 also showing a persistent significant loss of expression when compared to WT hTERT-1604s even more so than that in the UH4 cell line (Figure 5.6C).

To further investigate the role of UHRF1 and implicate the PHD domain in the maintenance of genomic imprints, I attempted to generate an HCT116/HT29 KO and mutant cell line similar to the HEK293K mentioned above. As HCT116 appears to be more tolerable than that of the HEK293K counterpart, HCT116 and HT29 seemed like good models to attempt these edits for a second time (Liao et al., 2015; Rhee et al., 2000). Unfortunately, the PHD mutant had no clones grow from a single cell that contained the mutation, two HCT29 KO clones did show healthy growth and were seen with the KO CRISPR mutation. However, both clones did not show any signs of a successful KO during pyrosequencing and qPCR investigations (Appendix Figure A2).

Figure 5.6

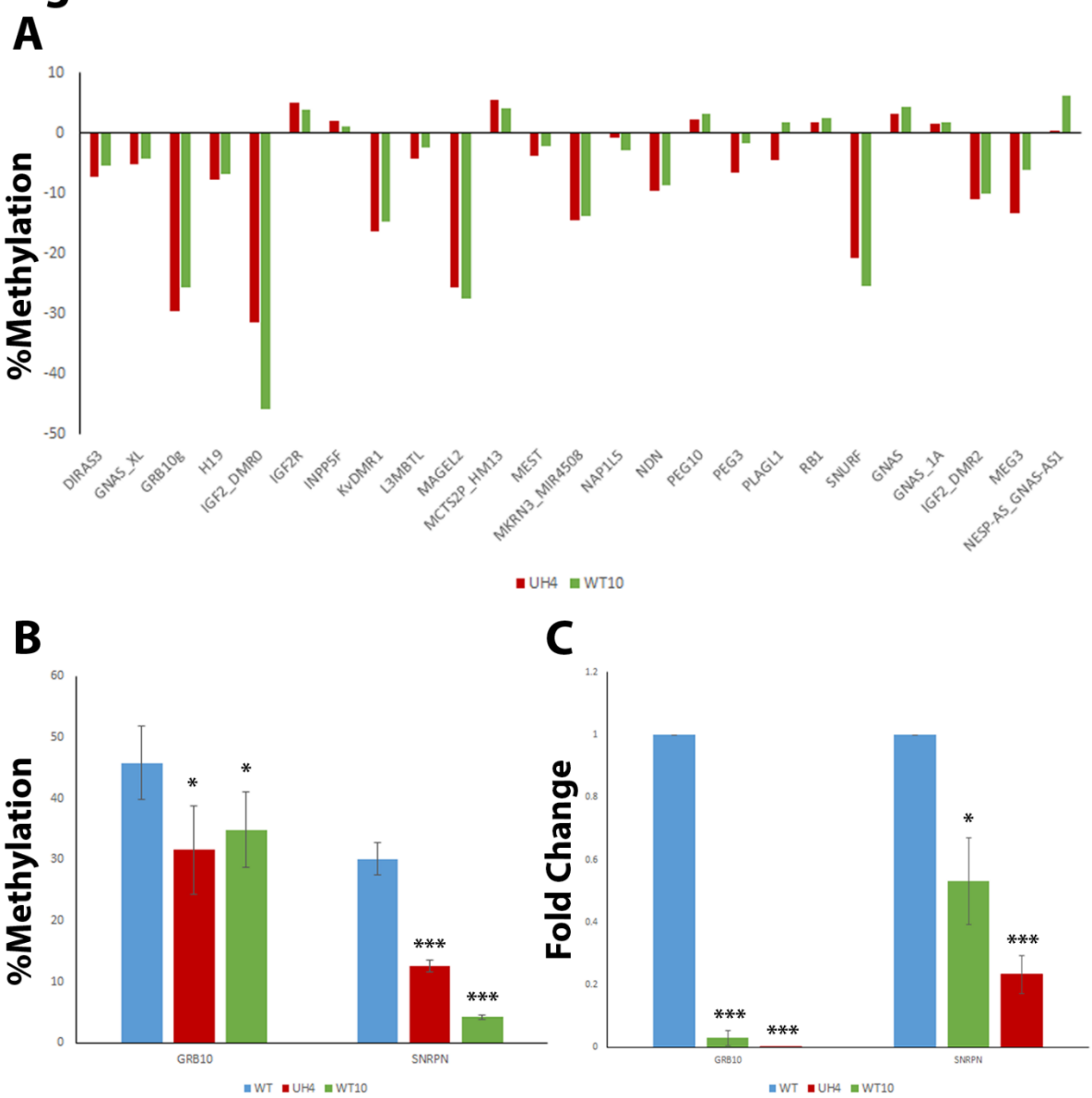


Figure 5.6: Depletion of UHRF1 in a differentiated human cell line results in loss of imprinting and failure to recover upon rescue. (A) Long term depletion of UHRF1 in the hTERT-1604 (UH4) cell results in loss of methylation at known imprinted DMRs as defined by Woodfine, Huddleston, and Murrell 2011 and Court et al. 2014 according to Infinium 450k beadchip array the UHRF1 rescue (WT10) failed to recover the loss of imprinting. **(B)** This loss of methylation was confirmed in *GRB10* and *SNRPN* in both the UH4 and WT10 cell lines using pyrosequencing. **(C)** The loss of methylation at the imprinted DMRs for *GRB10* and *SNRPN* resulted in the loss of expression for the corresponding loci, however, these changes were unexpected for SNRPN (Glenn et al., 1996). On the other hand, the imprinting of GRB10 is more convoluted in human and has not been fully described but methylation at this homologous *GRB10* DMR in mouse is linked with the repression of the gene (Hikichi et al., 2003). Pyrosequencing and qPCR experiments were carried out as duplicates with biological replicates. Error bars indicate standard error of the mean. Significance carried out with a two-tailed T-test with the following key, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

5.6 Discussion

Here we showed that the mutation of the PHD domain of *Uhrf1* in mouse embryos failed to gain methylation at the retrotransposable elements resulting in their subsequent activation, further contributing the current literature that suggests DNA methylation plays an important role for long term suppression of these elements. Further to this, we showed that in an adult human cell line that in a stable depletion of UHRF1 resulted in the derepression of the ERVs with repression upon the rescue of the WT UHRF1 protein. The PHD domain was then implicated further when the PHD mutant protein failed to rerepress the ERVs to the same extent of the WT protein. We then went on to be the first to show that a mutation at the PHD domain of Uhrf1 results in loss of imprinting at a subset of imprinted genes during embryonic development. Finally, we contribute to the current literature in regards to the recovery of imprinting in UHRF1 models by showing that UHRF1 and subsequent WT protein rescue does not result in the recovery of imprints in a chromosomally stable human adult fibroblast cell line.

UHRF1 plays an essential role in maintaining methylation patterns throughout the genome. Since abnormal methylation patterns have long been associated with cancer, UHRF1, for this reason, is currently being actively researched as a potential therapeutic target for cancer (Baylin and Jones, 2016). AZA developed by Píškala and Šorm 1964 is a demethylating agent used in cancer therapy that is known to cause genome-wide promoter DNA hypomethylation and subsequent gene re-expression and anti-tumour changes in key cellular regulatory pathways (Tsai et al., 2012). AZA acts as a demethylating agent by incorporating into replicating DNA and covalently trapping DNMT1 and indirectly inhibiting it (Yoo et al., 2007). As well as anti-tumour changes in key regulatory pathways AZA is known to activate a ‘viral mimicry’ response targeting cancer-initiating cells by causing dsRNA expression (Roulois et al., 2015). Recent investigations from the Walsh lab found that the UHRF1 depletion in the hTERT-1604 also resulted in a viral mimicry state which is not shared by depletion of DNMT1 in the hTERT-1604s. As mentioned in the introduction the Walsh lab previously carried out point mutation recovery experiments that implicated the PHD H3K9 binding, during this study we confirmed the stabilisation of the PHD mutation using a western blot. To further investigate the PHD domain of *UHRF1* and its implication in the repression of transposable

elements a PHD mutation was carried out in a mouse model using the same point mutation suggesting that the protein would still be stable. While the expression of Uhrf1 was normal it would be beneficial to confirm the stabilisation of the PHD mutant protein in the mouse model using a western blot. With the use of this model, it was possible to further show that the derepression of the retrotransposable element IAP occurs in mouse embryos at both e8.75 and e9.5 following mutations at the PHD domain of *Uhrf1*. Sequencing of the bisulphite converted DNA confirmed that the methylation lost was across several different *IAP* repeats. A knockout of *SETDB1* carried out by Karimi et al (2011) in mESCs showed that ERVs lose H3K9me3 and results in their subsequent de-repression. This may help explain the loss of methylation caused by the PHD mutation as this domain is essential for the binding of Uhrf1 to the H3K9me3 (Karagianni et al., 2008). Studies have shown a link between H3k9me3 and DNA methylation with some overlap being observed in mESCs (Karimi et al., 2011). However, in post-implantation, the correlation between DNA methylation and H3K9me3 is seen to be of greater importance (Wiznerowicz et al., 2007). Further studies have also noted that retrotransposon suppression in mice during preimplantation is mediated largely by SETDB1 with DNA methylation having a greater role in post-implantation (Ecco et al., 2016; Matsui et al., 2010; Wiznerowicz et al., 2007). Previous studies have also revealed that KAP1 is recruited by the KRAB-ZNF transcription factor family to the ERV elements and plays an important role for their suppression (Ecco et al., 2016; Tie et al., 2018; Wiznerowicz et al., 2007). Further recruitment of the DNMT's and SETDB1 have been observed to be partly dependent on KAP1 (Ecco et al., 2016; Wiznerowicz et al., 2007).

Similarly to the *IAP* elements, the methylated allele of ICRs is also normally associated with H3K9me3, which was observed to be selectively bound by ZFP57 and its cofactor KAP1 which co-immunoprecipitates with Uhrf1 (Quenneville et al., 2011). Therefore, this PHD mutation of *Uhrf1* in mouse embryos proved to be an interesting model for a novel investigation into the role of the PHD domain in the maintenance of genomic imprints during embryonic development. Uhrf1 has been observed in playing a role in genomic imprinting and retrotransposable by being involved with both the maintenance and the *de novo* methylation at the genomic imprints as reviewed by Unoki (2019). This group showed that the loss of Uhrf1 in fully grown oocytes only had a minuscule effect on the methylation at maternal ICRs except at *Gnas1A*, *Peg10* and *Mest* ICRs. However, while having a minuscule effect on the

methylation at the ICRs in oocytes, the authors observed that Uhrf1 was involved with 25% of genome-wide *de novo* methylation in oocytes (Maenohara et al., 2017). The KO of Dnmt1 in oocytes only results in a small loss of methylation genome-wide and Uhrf1's ability to interact with *de novo* Dnmt's it is possible that Uhrf1 interacts with Dnmt3a during *de novo* methylation in oocytes (Maenohara et al., 2017; Meilinger et al., 2009; Shirane et al., 2013). However, preimplantation embryos lacking maternal Uhrf1 lost over 3/4s of their methylation at all ICRs examined. Interestingly the loss of maternal Dnmt1 in blastocytes had much less severe effect on the ICRs with only 50% of methylation being lost at the ICRs. Furthering this Maenohara et al (2017) witnessed large losses of methylation at the *LINE1* and *IAP* elements in preimplantation embryos lacking maternal Uhf1. We furthered this work carried by implicating the PHD domain of *Uhrf1* for the maintenance of retrotransposable elements and genomic imprinting during later stages of embryonic development.

Furthering the *de novo* implication of Uhrf1 and recovery of methylation at a subset of gDMRs in *Uhrf1* null mice rescued with the WT protein similar to the method later used by us for *Dnmt3a2* (Qi et al., 2015; Thakur et al., 2016). Here we further explore the role of Uhrf1 in genomic imprinting by investigating the loss of imprinting in mouse embryos with PHD mutations at *Uhrf1*. While the loss of methylation is evident in the PHD mutation embryos the extent of the loss does not appear to be as severe as shown in mouse blastocysts lacking maternal UHRF1 in the study carried out by Maenohara et al (2017), suggesting that the complete loss of maternal UHRF1 is more severe to the maintenance of ICRs during early embryonic development. However, both the maternal knockout and PHD mutation of *Uhrf1* appear to have dramatic effects by presenting with partial preimplantation and embryonic lethality respectively (Maenohara et al., 2017). The worsened physiological effects observed in embryos at later stages could be due to the depletion of the maternal Uhrf1, as suggested in the case of DNMT1 mutants (Li et al., 1993): alternatively, the worsened abnormalities may also be caused by a domino effect of aberrant changes happening throughout the embryo developmental stages.

To further understand the role of UHRF1 in genomic imprinting in humans, we studied the imprints in a UHRF1 depleted differentiated cell line named UH4. The UH4 cell model only showed substantial losses of methylation at the *GRB10*, *SNURF, KvDMR, MKRN3_MIR4508*

and ZNF597_NAA60 DMRs. However, after rescuing the cells by transfecting a plasmid containing *UHRF1* there was no recovery of methylation in the UH4s, unlike the somatic restoration of imprints as seen in mouse ESC (Qi et al., 2015) suggesting it may not be possible in human. Secondly, the loss of methylation in the UH4s is not as widespread as what is observed in developmental models which may suggest that UHRF1 does not play as a vitally important role in maintaining imprinting status in differentiated cell models (Maenohara et al., 2017; Qi et al., 2015). While the depletion of UHRF1 appears to have a severe impact on the transposable elements as shown by Kong et al (2019), the imprints may require a complete knockout in human cell models to cause a widespread loss of imprinting.

The Walsh lab has previously attempted to generate two human cell model with (1) UHRF1 knocked out and (2) a mutation at the PHD domain using the HEK293T cell line. However, this appeared to have lethal effects on the cells and no positive cells could grow from a single cell. Continuing from this we decided to use human colorectal cancer cell lines HCT116 and HT29 as they have been previously reported to tolerate DNMT knock-outs which were observed to be lethal in hESCs suggesting they may have been a more suitable model for the *UHRF1* CRISPR edits (Liao et al., 2015; Rhee et al., 2000). However, after clonal expansion from mCherry positive single-cells separated using FACS, only two clonal HCT116 cell lines showed promise with single base-pair homozygotic frameshift mutations present within *UHRF1* resulting in altered expression of UHRF1. However, using *LINE1* as a guideline of global methylation, the frameshift KO cells unexpectedly exhibited no loss of methylation at these repeats. Further investigations into the expression of genes known to be controlled by methylation and which have also been previously shown to be disrupted in UHRF1 depleted cell models exhibited no change of expression between the suspected frameshift mutant HCT116 cell lines and the WT HCT116. Possible reasons for the false-positive effects observed in the HCT116 CRISPR edited cell lines could be that the *UHRF1* locus could be aneuploid, which can be a common issue in tumour cell lines, additionally, the single base frameshift mutation may not have been sufficient to induce a phenotypical mutation (Barrangou et al., 2015).

We showed here that the PHD domain of *UHRF1* plays a vital role in the maintenance of transposable elements and a small set of imprints during mouse embryonic development. It

would have been valuable to further investigate the methylation at other imprinted DMRs and the expression of their respective imprinted genes. However, due to working with mouse embryos, the quantity of material was a limiting factor. Proceeding it may be beneficial to generate an mESC cell line from the PHD mutant mice if possible. This would not only remove the sample material limitation it would allow for further in-depth analysis. One such analysis could be to use CRISPR-CAS9 to correct the mutation at the PHD domain in an attempt to recover methylation and expression of the transposable elements and imprints which has been observed respectively by the Walsh lab and Qi et al (2015) in similar model systems. In an attempt to generate a true *UHRF1* PHD mutant or KO in a human cell model, it may be beneficial to redesign the CRISPR-CAS9 sgRNA and plasmids and repeat the CRISPR to confirm that these edits are indeed lethal in these models. Further to this, it would be interesting to carry out other further *UHRF1* domain edits using CRISPR to understand fully which domains are lethal when mutated in human cell models and the different effect they may have at methylation-dependent elements.

5.7 Summary

This chapter demonstrates that the PHD domain of *Uhrf1* plays a role in maintaining retrotransposable elements and genomic imprinting during mouse embryonic development and appears to be highly lethal resulting in early embryonic termination around e9.5. Further to this, it could be seen that the depletion of UHRF1 causes loss of genomic imprinting at a small subset of imprints with failure to rescue upon recovery of UHRF1 in a human differentiated cell model. Finally, this chapter implies that mutations at the PHD domain of *UHRF1* and/or the complete knock-out of *UHRF1* is fatal to human developmental colorectal cancer cell models. It is clear that UHRF1 plays an important role in maintaining methylation at several different elements throughout embryonic development and in differentiated cell models, due to its implications shown here and in the literature it would be an interesting target for therapeutic interventions for diseases such as cancer and genomic imprinting disorders. However, further research will need to be carried out to establish such therapeutic interventions.

Chapter 6- General Discussion

6.1 Results Summary

In this thesis, we have shown crosstalk across gametic and somatic DMRs in a long-term knockdown of DNMT1 in the chromosomally stable hTERT-1604s. This crosstalk to a certain extent shows a level of reprogramming in which the loss of methylation at a subset of gDMR leads to the gain of methylation at a corresponding sDMR at the same imprinted cluster. This unique mechanism hinted at a potential DMR at the promoter region of NLRP2, which upon further investigation showed preferential maternal expression and a potential paternal origin of methylation at the germline. Further to this, we demonstrated a possible statistical method in which to score 'normal' methylated DMRs against 'abnormal' by creating a methylation index. This methylation index was able to accurately flag disturbed regions in previously diagnosed patients and even identified a patient that was previously diagnosed with SRS as a misdiagnosis. Finally, we furthered the current understanding of UHRF1 in the maintenance of retrotransposable elements and genomic imprints using a mouse developmental model and a human differentiated model. The mutant mouse embryos indicated that the PHD domain of Uhrf1 plays a vital role in the repression of the retrotransposable elements as well as maintaining normal methylation and expression at the imprints during embryonic development. Consistent with this, the PHD mutant was observed to result in early termination in homozygotic mutant embryos. Furthermore, an shRNA depletion of UHRF1 in a human fibroblast cell line resulted in the loss of imprinting at a small subset of imprints with failure to recover upon the rescue of UHRF1. Finally, the CRISPR KO and PHD mutant of *UHRF1* appears to be lethal in human cell models due to the failure of establishing a true KO or mutant *UHRF1* in human colorectal cancer cell lines.

6.2 Genomic Imprints

6.2.1 DMR Hierarchy

Somatic imprints are aptly named due to their unique establishment of the methylation mark after fertilisation. The mechanism for the *de novo* methylation at the sDMRs is still currently unknown, however, it is suggested that it may be guided by other epigenetic marks that were established during early development (John and Lefebvre, 2011). Further to this, the literature is heavily supportive that the establishment of methylation at the sDMR is very much dependent on the hierarchy interaction of the gDMR, as observed in a subset of imprinting clusters controlled by the gDMRs *H19*, *KvDMR*, *IG-DMR* and *GNAS* (Anwar et al., 2012; Elli et al., 2018; Kagami et al., 2010; Srivastava et al., 2003). Figure 6.1 summarises our current understanding of the hierarchy system found between some gDMRs and their respective sDMRs. In the DNMT1 depleted hTERT-1604 cell model we showed a 'methylation switch' form of reprogramming at a subset of imprints which appears to be unique to the long-term knockdown of DNMT1. John and Lefebvre (2011) suggested that investigations into sDMRs may provide insight into *de novo* methylation outside gametogenesis and may help shed some light on the complicated mechanisms that are involved in gene silencing during embryonic development and differentiation.

As mentioned sDMRs are relatively rare and correlate with gene promoters and transcription factor binding sites, higher-order chromatin structure, chromatin modifications and lncRNAs (Barlow and Bartolomei, 2014; Monk et al., 2019). It has been suggested that *de novo* enzymes targets sDMRs from epigenetic marks established in the germline, it has recently been proposed that H3K27me3 is responsible for this (Hanna et al., 2019). Once established the sDMRs are maintained by DNMT1, in which DNMT1 is recruited to sDMR's through an RNA-dependent mechanism to maintain the silencing of ubiquitously imprinted genes (Kanduri, 2016). This is interesting as it would be expected that the depletion of DNMT1 would result in the loss of this methylation, but using the hTERT-1604 DNMT1 and UHRF1 depleted models only the depletion of UHRF1 results in widespread loss of methylation at the sDMR while the DNMT1 depleted model either gains or maintains methylation at some sDMRs. The gains of methylation may be independent of the gDMR in this model and rather in the long-

term depletion of DNMT1, this RNA dependent mechanism is still functionally active recruiting the remaining DNMT1 protein. Alternatively, the long-term loss of methylation originally caused by the DNMT1 depletion may result in epigenetic compensation such as histone modifications at the sDMRs re-establishing the mark to recruit *de novo* methyltransferases to the sDMRs (Barlow and Bartolomei, 2014). To investigate potential changes of epigenetic regulators such as histone modifications, binding of transcriptional factors, and continued binding of DNMT1 or recruitment of *de novo* enzymes at the sDMRs, would have to be investigated using tools like ChIP-Seq. Further to this to confirm whether or not the changes at the sDMR is dependent on the gDMR or epigenetic compensation would require further experimentation. However, understanding this mechanism may shed light into the complicated mechanism in gene silencing in embryonic development and differentiation.

Figure 6.1

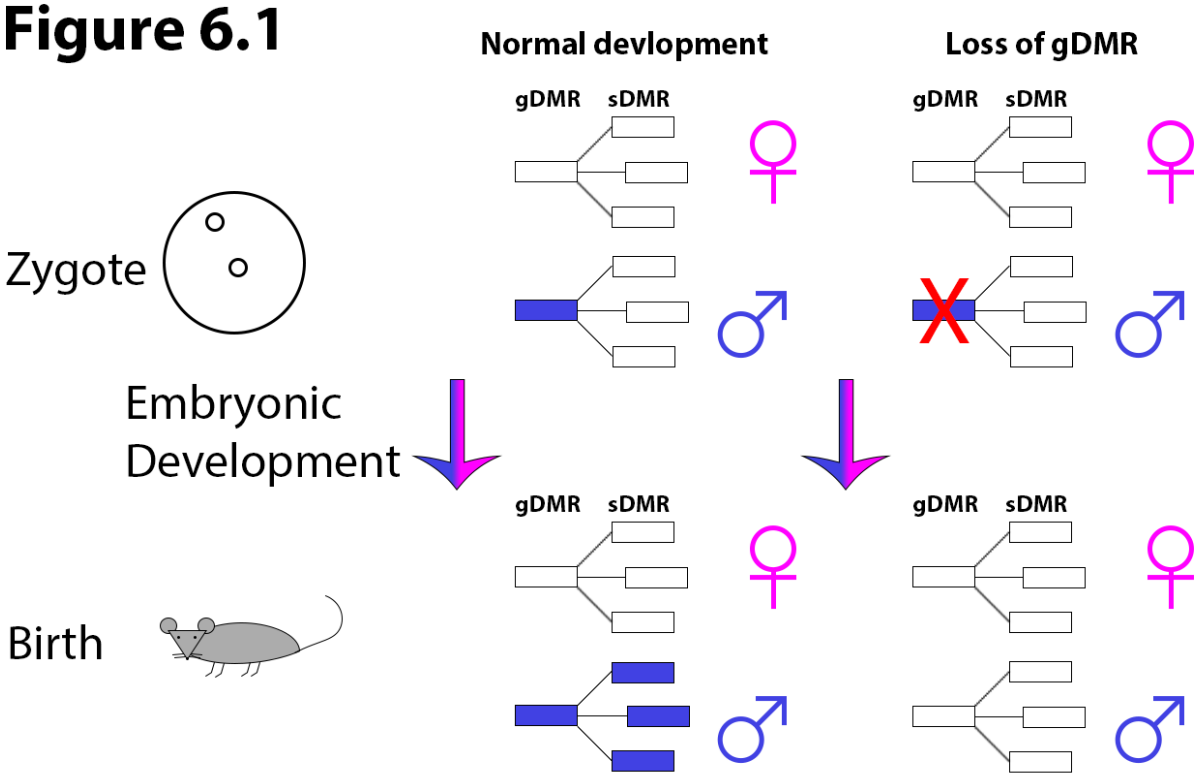


Figure 6.1: Schematic diagram describing the hierarchy of control of methylation at sDMR by the gDMR. As shown previously in a figure earlier in the thesis, the methylation at the gDMR is established in the germline, while the sDMR is established later during embryonic development. However, in this figure, it can be seen that one gDMR contains a hierarchal function over several sDMRs, with both the gDMR and sDMR being methylated on the paternal allele after birth. However, the loss of the gDMR before the establishment of methylation at the sDMR can lead to the loss of parental-specific methylation at the sDMRs due to the hierarchical function carried out by the gDMR.

Cross talk among primary and secondary DMRs has previously been associated in a number of patients with imprinting disorders with MLID: a study by Court *et al* (2013) observed changes of methylation at the gametic *NESP-DMR* were accompanied with loss of methylation at the nearby somatic *GNAS-XL* and *GNAS-EX1A* DMRs. We observed a similar hierarchal interaction between the *H19* gDMR and the nearby sDMR *IGF2_DMR0* in which all BWS patients with hypermethylation at the *H19* gDMR were also observed to have hypermethylation at the *IGF2_DMR0*. Similarly, almost all SRS patients with hypomethylation at the *H19* gDMR were observed with hypomethylation at the *IGF2_DMR0*. Moving forward it may be reasonable to suggest that altered methylation at the *IGF2_DMR0* is a common disturbance in SRS and BWS patients and should be investigated using the normal MI. This interaction has been described previously in other studies in which the *IGF2_DMR0* and *IGF2_DMR2* were both seen gaining methylation along with the gametic *H19* DMR in BWS patients and it has been suggested that the methylation for both *IGF2* DMRs were established at the same time (Maeda et al., 2014). Further confirming this interaction, it can be observed that the *IGF2_DMR2* and *IGF2_DMR0* are both hypomethylated in the DNMT1-depleted hTERT-1604s, following the same direction of loss of methylation at the gametic H19 DMR. While the *IGF2_DMR2* has been associated with expression of Igf2 in mice, the function of *IGF2_DMR0* remains unknown (Maeda et al., 2014; Murrell et al., 2001).

Maintenance methylation and imprints

In recent years it has been clear that DNMT1 is not sufficient for the maintenance of DNA methylation and requires all aspects of the PCNA/DNMT1/UHRF1 complex to faithfully maintain methylation, any disruptions to this complex has been observed to result in cancer-specific methylation (Pacaud et al., 2014). Further to this interference with the UHRF1 associated complex such as Histone deacetylase 1 (HDAC1), G9a and GLP also result in cancer-specific methylation (Arzenani et al., 2011; Dong et al., 2008). As mentioned in the earlier chapters of this thesis UHRF1 plays a vital role in co-localizing DNMT1 to the hemimethylated DNA, so it is not a surprise that the interference with UHRF1 or its associated complex results in loss of maintenance methylation. However, while DNMT1 and UHRF1 work together to maintain methylation and result in a similar global loss of methylation when either protein is knocked out, UHRF1 appears to have a more severe impact at specific regulatory regions.

From previous work in the Walsh lab, it was observed that the depletion of UHRF1 in hTERT-1604s resulted in a similar global loss of methylation in comparison to the DNMT1 depletion in the same cell model. However, the UHRF1 depletion resulted in a transcriptional state of 'viral mimicry' due to a much greater upregulation at a few retroviral elements such as Interferon-Stimulated Genes (ISGs) than was observed in the DNMT1-depleted cell model. Similarly, the well-investigated imprinted DMRs such as *GRB10*, *SNURF*/*SNRPN*, *KvDMR* and *H19* also appear to be more hypomethylated in the UHRF1 depleted hTERT-1604s than in their DNMT1 counterpart. As previously discussed this similar methylation response at the imprinted ICRs was reported in Uhrf1 null mouse embryos which were observed as having greater loss of methylation than was seen in Dnmt1 null mouse embryos (Maenohara et al., 2017).

6.2.2 Is the NLRP2 promoter region a novel DMR?

The unique methylation pattern at the DMRs in the DNMT1-depleted hTERT-1604 model flagged the imprinted gene *NLRP2* as having increased methylation at a previously described DMR and loss of methylation at the promoter region (Woodfine et al., 2011). As the promoter was not described previously as a DMR, further investigations were carried out at the *NLRP2* promoter region to verify its potential DMR status. While this investigation uncovered paternal specific methylation and maternal expression in aESC and phESCs, there were no other signs that this region was imprinted. A study by Baran *et al* (2015) claimed that *NLRP2* was a false positive due to picking up paternal and maternal monoallelic expression in their initial validation of imprinted genes in whole blood in the same tissues but across different samples, which I have termed as 'unfaithful'. However, while Baran *et al* (2015) decided to write *NLRP2* off as a false positive, it seems that the potential imprinting status of NLRP2 is more convoluted and potentially interesting than once believed. While not shown in the main body due to its conflicting nature with the parental ESC models, the promoter region of *NLRP2* in genome-wide uniparental disomic patients suggested paternal specific methylation. However, there was though no RNA-seq available to confirm if the expression had also switched with the parental specific methylation.

It is hard to confirm whether this novel DMR present at the promoter region of NLRP2 is established in the germline or somatically without further testing. In support of the DMR being established in the germ-line, my study showed preferential methylation and expression bias in aESCs and phESCs which are believed to mimic levels of methylation seen in sperm and oocyte (Kim et al., 2007; Revazova et al., 2007). Furthering this, a study by Vértesy *et al* (2018) showed a higher expression in female germ cells than that observed in male, which would suggest that the promoter of *NLRP2* in the male germ cells would contain higher levels of methylation and potential gametic establishment of the imprinted mark. However, this maternal specific expression only appears to persist into the early stages of pregnancy (Pilvar et al., 2019).

6.2.3 non-canonical establishment of NLRP2 in the placenta

A study carried out by Sanchez-Delgado *et al in* 2016 described a mechanism in which oocyte methylation survived past the blastocyst stage and persisted as transient DMRs. This would explain why Pilvar *et al* (2019) later detected monoallelic expression mostly during early-stage pregnancy but does not explain the 'unfaithful' monoallelic expression observed in somatic tissues (Baran et al., 2015). Interestingly *NLRP2* appears to have a unique mechanism in which it can switch from biallelic expression to monoallelic expression during cellular reprogramming (Jeffries et al., 2016). With this in mind if the monoallelic expression is lost after early pregnancy, later cellular differentiation may lead to reprogramming of a somatic DMR indiscriminately on one allele which would help explain the 'unfaithful' monoallelic expression described by Baran *et al* (2015). A similar model has recently been claimed to explain the establishment of a subset of sDMRs in which oocyte-derived H3K27me3 resulted in non-canonical and methylation-independent transient imprinting during preimplantation development (Figure 6.2). However, in the post-implantation extra-embryonic lineages a differential methylation mark I.E an sDMR is established as seen in Figure 6.2 (Hanna et al., 2019). Figure 6.2 contains a graphical abstract that explains the loss of H3K27me3 and the gain of biallelic methylation or the maternal methylation in the embryonic and extra-embryonic lineages respectively. Interestingly, H3K27me3 is lost on both alleles and methylation is the main repressor at these stages, however, the active H3K4me3 histone mark is still present on the active allele (Figure 6.2). In human, during the prezygotic genome

activation, widespread depletion of H3K27me3 occurs, suggesting that a similar imprinting model may be occurring in human as observed in mouse (Xia et al., 2019). Supporting this type of model for the novel NLRP2 DMR is an expression analysis study which showed a gradual decrease of expression in NLRP2 after fertilization and a gradual increase at day 5 (Zhang et al., 2008). However, Hanna *et al* (2019) observed that these non-canonical imprints were localised to the retroviral elements such as endogenous retrovirus-K (*ERV-K*) and long terminal repeats which is not seen at this novel *NLRP2* DMR, but a DNA repeat is present. Another possible explanation for its distorted imprinting status could be that it is polymorphic which may be due to an SNP leading to parental specific methylation pattern based on the parent it is inherited from. Whether or not these suggestions are plausible, or the region contains a mechanism not yet explored, further investigations are surely required if we are to fully grasp what is occurring at this locus.

Figure 6.2

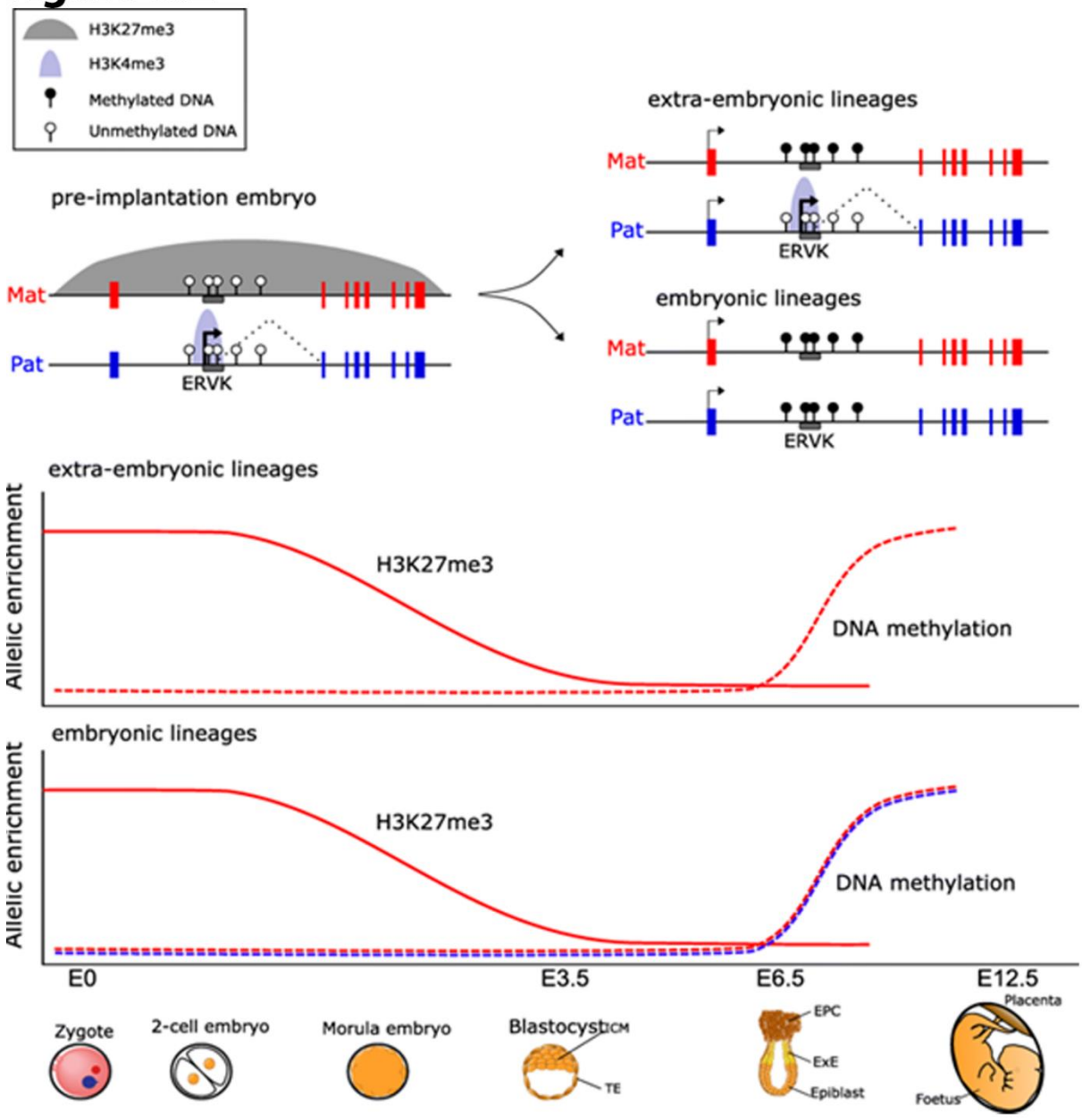


Figure 6.2 Noncanonical establishment of somatic differentially methylated regions in early development. In the preimplantation embryo, the sDMR containing the retroviral element *ERVK* is hypomethylated and is expressed in a monoallelic fashion from the paternal allele with the repressed maternal allele being observed with enrichment of the repressive histone mark H3K27me3 while the paternal allele is enriched with the active H3K4me3 mark. However, the maternal H3K27me3 mark is gradually lost throughout preimplantation development and is completely lost at post-implantation. However, in extra-embryonic lineages methylation is established only on the maternal allele resulting in repression of *ERVK*, while the paternal allele shows monoallelic expression of *ERVK* and is hypomethylated with H3K4me3 histone enrichment. In the embryonic lineages, DNA methylation is established on both the maternal and paternal alleles and loss of the H3K4me3 mark resulting in the repression of the *ERVK* element. ("http://creativecommons.org/licenses/by/4.0/") Reprinted from Hanna *et al.*, (2019).

6.2.4 Imprinting disorders and diagnosis

As the use of genome-wide technology increases in the investigation of imprinting disorders so does our understanding. One study investigating the prevalence of imprinting disorders in Estonia between 1998-2016 observed an increased prevalence of all imprinting disorders over this period, with an increased incidence of rare imprinting disorders in recent years which the authors suggest is due to advances in diagnostics and physician awareness (Yakoreva et al., 2019). Similarly, as research laboratories increasingly use genome-wide technologies during their investigations into imprinting disorders it is starting to become clear that there is a growing number of patients with imprinting disorders that suffer from MLID, rather their diagnosis being limited to one disease-specific locus (Eggermann et al., 2014; Eggermann et al., 2015c). One current issue with the investigation of DMRs as described by Monk *et al* (2018) is the definition of what is loss or gain of methylation. Along with this, the authors stated many points that would need to be answered which are (1) the statistical cut off (2) a number of controls required to define the 'normal' range (3) complications with mosaic epimutations (4) methylation variance at the CpGs within the DMR and (5) the reproducibility of the molecular techniques used. When investigating the possibility of using the Infinium bead-chip array as a tool for diagnosis of imprinting disorders, in doing so we defined a statistical cut-off for a normal range of methylation for a large set of DMRs. The MI defined in chapter 4 of this thesis for a set of 'normally' defined imprints in 100 samples answers the issues set by Monk *et al* in 2018 and was able to accurately score epimutations in previously diagnosed imprinting disorder patients as well as describing a potential misdiagnosis, presented in Figure 6.3. The MI I have described is dependent on the normal variability of methylation at each individual DMR across 100 patients, this means that with this MI would identify abnormal methylation that was completely dependent on the DMR. As an example, in Figure 6.3 has a normal variability between ~54% and ~63% at the *H19* DMR, if a patient were presenting with *H19* DMR methylation at 40% would be classed as abnormal using this MI. However, using the general characteristic of 35-65% this would still be within 'normal methylation' for an imprinted DMR. This is important as this would be a decrease of ~13% in methylation when compared to the normal blood sample with the lowest methylation value at the *H19* DMR which is quite a steep decrease. It is also worth to note that another study has also previously described patient methylation values of about 40% in SRS patients,

supporting the fact that the DMR though within the normal characteristic range may still be abnormally methylated (Gicquel et al., 2005). However, on the other hand 40% would be classed as normal methylation for the *PEG10* DMR as seen in Figure 6.3. It is easy to generalise imprinted regions with a generic methylation range of 35-65% due to the monoallelic methylation mechanism that they possess, while this range is suitable for initial characterisation of imprinted regions it is imperative that each region is investigated as individual regions.

Figure 6.3

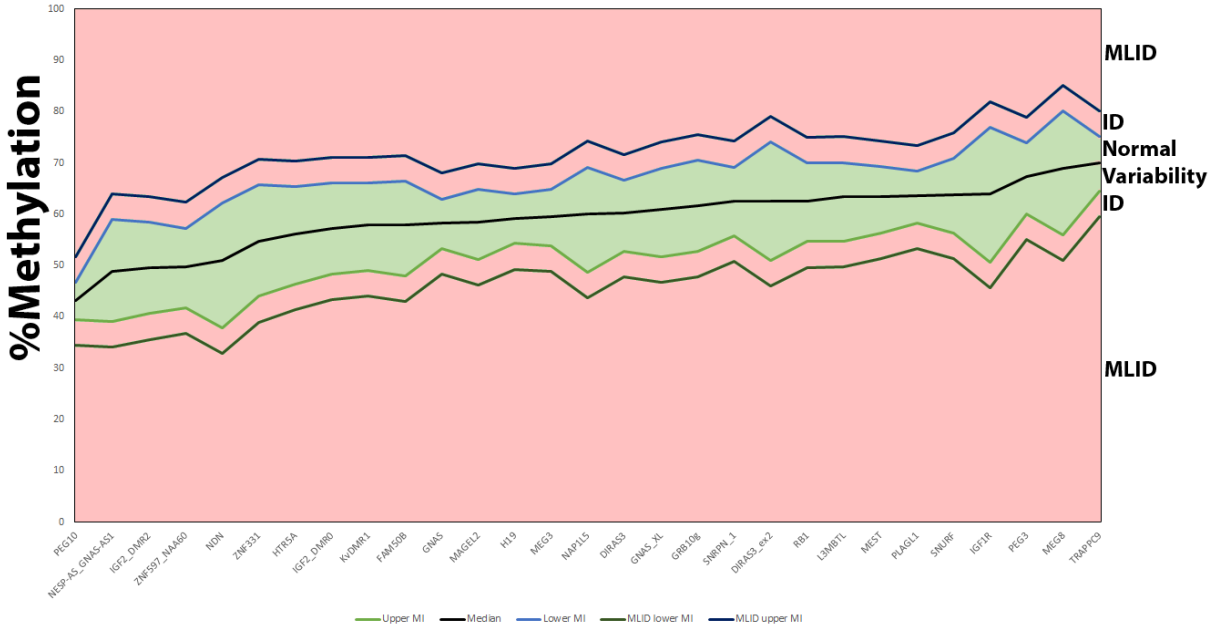


Figure 6.3: Graphical representation of the methylation index across select DMRs in blood. Using 100 patient blood samples the average methylation for each DMR was examined for viability. As the methylation at the DMRs across the samples was not normally distributed we settled on using the median and IQR to create a methylation index (MI). Using the 1.5 x IQR rule to find outliers we created the upper and lower limits which was applied to the commonly disrupted DMRs linked with imprinting disorders. The methylation inside the MI for each region highlighted green on the figure would be classed as normal variability, while outside the upper and lower MI highlighted red on the figure and labelled with ID (Imprinting disorders) would be classed as an outlier of normal variability for the region and therefore can be used to identify abnormally methylated DMRs. The MI was extended further with an extra + or -5% added to the upper and lower limit respectively to detect potential multiple-locus imprinting disturbances (MLID). This more stringent limit was implemented as these regions are not fully described in their functionality or implication within imprinting disorders and therefore we wanted to only identify more 'extreme' changes of methylation to increase the likelihood that the region is implicated with the imprinting disorder being investigated.

Imprinting disorders are usually caused by aberrant methylation patterns present at the ICR of the imprinted loci. However, a sub-group of imprinting disorder patients are observed with epimutations at other loci which have been discussed as potentially altering gene expression at multiple imprinted loci, resulting in additional clinical features (Begemann et al., 2018). Due to these complications with MLID, research must focus on understanding the physiological effects that these epimutations outside the disease-causing loci cause and the mechanism for how they occur. Multiple studies have called for an urgent need for multi-locus imprinting disorder testing and improved diagnostics in general for imprinting disorders, with the bead chip array being discussed as a potential future diagnostic tool (Eggermann et al., 2014; Eggermann et al., 2016; Grafodatskaya et al., 2016).

6.2.5 Multi-locus imprinting disorders

As genome-wide analyses of imprinting disorders become more prominent patients with imprinting disorders are presenting with additional disturbances at imprinted regions outside the disease-causing loci, however, it is not known if and to what extent which MLID may or may not be affecting the phenotype of the imprinting disorder (Sanchez-Delgado et al., 2016a). Maternal effect mutations in members of the subcortical maternal complex (SCMC) such as the PADI6, NLRP5 and NLRP2 have been observed in MLID (Demond et al., 2019; Meyer et al., 2009; Qian et al., 2018; Sparago et al., 2019). The SCMC in oocytes and preimplantation embryos has long been speculated to be involved in the establishment and maintenance of maternal imprints during early development. Maternal mutations in many of the members that make up the SCMC have been observed to cause loss of imprinting and or early embryonic termination. However, further investigations are required to fully understand the role of the SCMC in genomic imprinting (Monk et al., 2017). Interestingly as NLRP2 is one of the components of this complex it could help further support a role for the novel DMR in which the non-canonical imprinting mark being established at post-implantation extra-embryonic ectoderm stage helps to control the gene dosage of *NLRP2* once it has completed its major role in the SCMC.

As mentioned NLRP2 has been associated with MLID: one mother contained a homozygous (Arg493SerfsTer32) variant of *NLRP2* and had suffered 3 miscarriages and gave birth to two children, both afflicted with BWS and MLID. Both children were heterozygotes for the *NLRP2* variant (Begemann et al., 2018). Within the same study the authors discussed a mother with a predicted deleterious mutation in the TTD domain of UHRF1 (Val172Met), this mother gave birth to discordant monozygotic twins both with the *UHRF1* variant. However, only one suffered from SRS with MLID while the co-twin was epigenetically normal.

6.2.6 Retroviral elements and imprinting

In the previous chapter, the role of the PHD domain of Uhrf1 was investigated in the repression of the retroviral element *IAP* in mouse embryos. This investigation along with previous investigations into Uhrf1 carried out by the Walsh lab agrees with much of the literature which describes a clear role for UHRF1 in the repression of transposable elements (Chernyavskaya et al., 2017; Kong et al., 2019; Sharif et al., 2007). The experiments carried out previously by the Walsh lab in the hTERT-1604 UH4 and recovery models, along with the results from the experiments carried out in this thesis, heavily implicates the H3K9 binding function by the PHD domain of UHRF1 in the repression of transposable elements and silencing of the inactive allele at genomic imprinted regions. However, one study by Sharif *et al* 2016 contradicts the implication of UHRF1 in the repression of the transposable elements, suggesting that in the absence of Dnmt1, Uhrf1 actually blocks Setdb1 from binding to the transposable elements and therefore blocks methylation-independent repression of the elements in mouse embryos. In support of Sharif *et al* 2016, a mouse liver regeneration model with a hepatic KO of *Uhrf1* resulted in even further repression of IAP-d-int due to epigenetic compensation of H3K27me3 from active promoter regions to the transposable elements in Uhrf1HepKO mice (Wang et al., 2019). Wang *et al* (2019) carried this out by using flox sites flanking exons 6–10 which when recombined resulted in a stop codon, deleting the nuclear localization signal and all the domains required for DNA methylation. These mice were then crossed with mice expressing the Cre recombinase under the liver-specific albumin promoter backcrossed onto the *Uhrf1fl*/fl background to generate Uhrf1HepKO mice. The deletion of *Uhrf1* resulted in the loss of methylation at the TE's and epigenetic compensation of H3K27me3 from the nearby promoter regions to the TE's after partial hepatectomy, resulting

in further repression of *IAP*-d-int in comparison to quiescent Uhrf1hepKO Liver. However, while being the direct opposite that is observed in studies carried out by the Walsh lab and others this could be explained due to the difference in the experimental design of these studies. For example, one major factor which would not be affected in these two studies but would in our model system is maternal Uhrf1 in the mouse embryos, which plays a vital role in maintaining methylation in preimplantation embryos (Maenohara et al., 2017). A recent study has shown that Uhrf1 is important for the repression of retrotransposons in male germ cells in mice and that the cKO of *Uhrf1* in germ cells also resulted in the upregulation of the imprinted loci *Dlk1* and *Gtl2* (Dong et al., 2019).

It was theorised by Denise Barlow in 1993 that the development of imprinting may have been due to the host defence system, a theory that has held up for a subset of imprinted loci as recently discussed by Ondičová, Oakey and Walsh, 2020. A recent study showed that parental-specific methylation at 17 human-specific imprinted gDMRs is a consequence of transcription of the loci initiating in promoters of nearby lineage-specific endogenous retroviruses (ERVs) leading to *de novo* methylation in the oocyte (Bogutz et al., 2019). As briefly discussed above, retroviral elements have been observed which overlap with imprinting transcription start sites that have inherited oocyte H3K27me3. This histone mark was then gradually lost in the preimplantation embryo along with the rest of the genomic H3K27me3, then later methylated maternally in extraembryonic lineages, establishing sDMRs for the investigated imprinted regions (Hanna et al., 2019; Zheng et al., 2016b). Hanna *et al* (2019) later reviewed unexplained factors from their study which was that the (1) associated imprinted gene expression of the ERV elements was regulated? by the methylation, (2) why sDMRs were established at the regions instead of the maintenance or re-establishment of the H3K27me3 and (3) the possible existence of unknown factors involved with either marking the inactive allele for methylation or protecting the active allele from methylation (Hanna, 2020). From recent studies, a subset of imprints has a clear overlap with retroviral elements and would prove to be an interesting area for further investigations.

6.3 Future direction

6.3.1 Further characterization of the beadchip array and diagnosis of imprinting disorders

Through the online bioinformatic software USEgalaxy, we have described a very bioinformatic friendly method to score imprinted DMRs in blood using the highly reproducible 450k array beadchip array, matched with the bespoke workflow designed by Thursby *et al* (2020). We have already made progress to investigate the more updated 850k epic array and seen a similar trend for scoring the imprinted DMRs in saliva. However, this was with a small sample size of 20 patients and may not be statistically representative as would be required for this type of work. Firstly, it would be paramount to greatly increase the sample numbers for the saliva investigation to get a large representation to ensure that the variability at each DMR is accurately accounted for. The reason for doing this in saliva over blood would be to create a non-invasive method to accurately diagnose imprinting disorders and fully investigate any underlying epimutations that may also have present as part of their disorder. Firstly, an accurate MI will have to be developed as described above. Once this has been completed and the DMRs and the defined MI is within the expected ranges it would have to be tested for its ability to identify epimutations in patients suffering from imprinting disorders. To do this, we would have to recruit a number of patients that have been previously diagnosed with an imprinting disorder and an epimutation at the disease-causing loci. The recruitment could also include a small subset of patients that have been diagnosed based on symptoms with no underlying epimutation at the known disease-causing loci to see if the saliva samples when analysed using the 850k epic array can identify epimutations in patients previously molecularly undiagnosed similar to the SRS study carried out by Prickett *et al* (2015).

While being a powerful tool for the investigations of imprinting disorders and the potential use as a diagnostic tool for imprinting disorders as shown in this thesis. The increase used of the 850k beadchip array for diagnostic, would not only provide a reproducible and informative diagnostic test hopefully reducing the risk of misdiagnosis of imprinting patients due to overlapping features. It will also open up opportunities to further investigate the genome-wide impact at the methylation level and increase our current knowledge in patients with imprinting disorders and the mechanisms behind the disorders which may in the future

lead to more effective treatment options. One aspect of imprinting disorders which will benefit greatly from a genome-wide methylation tool for diagnostics would be MLIDs if used. This would help to accurately represent the percentage of imprinting disorder patients that suffer from multiple epimutations, further allowing for researchers and clinicians to investigate and treat the patient based on the true extent of their methylome rather than just the disease-causing loci.

6.3.2 Reprogramming of methylation at imprints

As the recent research in the Walsh lab and the literature have both confirmed recovery of methylation at imprinted DMRs in mESCs models with the knockout and rescue of Dnmt3a2 and Uhrf1 it is clear that under the right conditions methylation at the imprints are susceptible to imprinting outside the germline (Qi et al., 2015; Thakur et al., 2016). While UHRF1 KO and rescue did not have the same effects in our human cells and it was subsequently not explored in this thesis further, investigations into the recovery of methylation in the Dnmt3a2 rescue cell lines would provide vital information for understanding reprogramming of imprints in somatic tissues. Qi *et al* (2015) showed that there was no correlation between the recovery of methylation and common histone marks. However, as clearly shown in this thesis the knockout of *Uhrf1* and *Dnmts* can have highly variable effects at the same loci. Therefore, this would be an opportunity for further exploration to confirm if histone marks are implicated in the recovery of methylation in our model system. Further to this, it would also be advantageous to rescue the *Dnmt3ab* KO mESC with clinically relevant point mutations in different domains of *Dnmt3a2*, this would help confirm what regions are required for the recovery of methylation at the imprinted DMRs. Keeping on the same direction if possible we could generate an mESC cell line from the Uhrf1 PHD mutant mice and rescue the mutant *Uhrf1* and investigate whether the loss of imprinting can be recovered as observed in the article by Qi *et al* (2015). A further novel investigation that if successful could provide exciting data would be to attempt to differentiate the Dnmt3a2 rescued mESCs into a neurological cell line to see (1) if it is possible, (2) would the neuralization process recover methylation at DMRs that failed to gain methylation and (3) will the imprinted loci at the DMRs that did recover methylation but showed insignificant changes of expression between the KO and rescue cell line, show significant changes of expression after neuralization?

While the novel switch of methylation between gametic and somatic DMRs in a human cell model with a long-term depletion of DNMT1 does not entail recovery of methylation at DMRs, it does appear to be a novel model describing possible cross-talk between DMRs within the same imprinted cluster. As already mentioned, wet lab investigations should be carried out at other imprinted loci that show the unique trend observed in this model. The model should also be further explored by examining the gametic and somatic DMRs for changes in histone modifications that may be occurring with the loss and gains of methylation at these DMRs. As this appears to be unique to the long-term depletion of DNMT1 in the hTERT-1604 these explorations into the histone modifications should also be carried out for the long-term depletion of UHRF1 and transient depletion of DNMT1 hTERT-1604 models for direct comparisons in an attempt to explain the differences in the models.

6.3.3 Human model for imprinting

While the hTERT-1604 cell line had benefits over other cell models used for human, such as being chromosomally stable etc, it is still restrictive in terms of being an accurate model due to experimental immortality. While the unique mechanism of gains of methylation at the sDMR is an exciting investigation observed in the DNMT1 depleted model, the hTERT-1604s do not appear to show expected changes of expression when methylation is lost at the DMRs in either the UHRF1 or DNMT1 depleted models, causing an obvious concern in terms of applying it to human imprinting disorders. Further to this more investigations will be required to fully understand the mechanism occurring at these sDMRs, for example confirming were the remaining DNMT1 protein in the depleted models are binding throughout the genome, this could be carried out using CHIP-seq analysis. Further to this CHIP-seq analysis may be carried out on the *de-novo* methyltransferases to see if *de novo* methyltransferases are responsible for the gains of methylation at the sDMRs. The Walsh lab has attempted to create UHRF1 KO/PHD mutant cell model in human ESCs and then in human colorectal cancer cells, however, it appears that these genomic edits by CRISPR in both cases appeared to have a lethal impact to the cells. A more suitable model would allow for more clinically relevant experiments to be carried out and help further the understanding behind imprinting and imprinting-like disorders in human. One way to carry this out would be to use primary blood samples and carry out DNMT1 and UHRF1 transient depletion with the use of a siRNA and

examine how the imprints respond to the siRNA treatment and how they respond after several days once treatment has stopped. A further model which would not have time restrictions which would be imposed by the primary cell models would be to use human ESCs. This would allow for long term depletions to be carried out with an shRNA if tolerated, like that used in the hTERT-1604 models. If the depletion does not appear to be lethal this would be a good model to see if the long-term depletion of DNMT1 mimics the methylation pattern observed at the DMRs which was observed in the hTERT-1604 model. Further to this, a UHRF1 depletion and rescue, if tolerated, would be a further model to investigate if recovery of methylation at the imprints could occur outside the germline in a chromosomally stable model, specifically in human.

6.4.4 Characterizing the novel NLRP2 DMR

The novel *NLRP2* DMR, if indeed it is a true imprinted DMR, seems to have very complicated regulation, but it may have a similar mechanism of establishment as described by Hanna *et al* (2019), due to role NLRP2 plays in the SCMC and its expression pattern in the early embryo. As of now this thesis only describes a methylation bias at the *NLRP2* promoter region using aESC and phESCs along with the unclear monoallelic expression described throughout the literature. However, while we indicate potential explanations, it would require intensive research to confirm these claims. Unfortunately, while *NLRP2* is present in mouse it does not appear to be imprinted meaning no mouse models could be utilised (Kuchmiy et al., 2016). The promoter region of NLRP2 does not contain H3K27me3 in hESCs according to the UCSC histone modifications track, which is aligned with the model that this histone mark is lost and replaced with methylation as an sDMR in the post-implantation tissue (Hanna et al., 2019). Xia *et al* (2019) showed a pattern of transcription and asymmetric H3K27me3 marking occurring after fertilization in human. Using the extensive dataset published on the Gene Expression Omnibus (GEO) with accession number GSE124718, it would be possible to confirm if maternal H3k27me3 is inherited from oocyte after fertilization and is then slowly depleted at the novel DMR. This can then be matched to methylation at each stage using the dataset under GSE81233 which contains single-cell methylome data for both male and female gametes and methylome data from the 2-cell to post-implantation stage. Finally, using the following dataset GSE36552 to comprehensively investigate the change of expression using

single-cell RNA-seq from oocyte until late-stage blastocyst, one could link *NLRP2* expression to any epigenetic changes observed in the previous datasets.

6.4 Concluding remarks

Using both wet-lab and bioinformatic techniques, I was able to investigate the loss of maintenance methylation of enzymes in normal adult cells as well as in mouse embryos in regards to genomic imprinting with the latter model also being used to investigate genomic retroviruses. My results add to the current knowledge of crosstalk among DMRs which can be observed at imprinting domains regulated by the hierarchical model. Further to this, I have described a methylation index to score normal methylation levels at several imprinted DMRs which can identify abnormal methylation at imprinted DMRs that are normally disturbed in imprinting disorders. Furthermore, here I showed that the PHD domain of Uhrf1 plays an important role in the regulation of maintenance methylation at genomic imprints in mice.

Appendix

Table A1

Imprinted DMR	Est	MO	Chr	Start	End	Blood	TS	MI Upper	MI Lower
KvDMR1	G	M	chr11	2719948	2722259	57.98	27.91	66.04	49.04
IGF2_DMR2	S	P	chr11	2153834	2155112	49.52	42.08	58.39	40.61
IGF2_DMR0	S	P	chr11	2168333	2170145	57.18	53.46	66.15	48.39
H19	G	P	chr11	2018812	2024740	59.20	55.83	63.95	54.29
MEG3	S	P	chr14	101290524	101293978	59.58	18.20	64.79	53.90
MEG8	G	M	chr14	101370741	101371419	68.89	84.56	80.09	56.03
NDN	G	M	chr15	23931451	23932759	50.99	52.52	62.15	37.88

Table A1 showing methylation at chromosome 14 imprinted DMRs in a Temples syndrome patient sample similar to the misdiagnosed SRS patient. Methylation across imprinted DMRs associated with the respective diseases, methylation values with red underlining text showing gains of methylation outside the defined methylation index (MI) while blue underlined text shows loss of methylation outside the defined MI. The patient shows clear hypermethylation at the *MEG8* DMR and hypomethylation at the *MEG3* DMR, suggesting possibly mUPD14, there is also significant hypomethylation at the *KvDMR* which was also observed in the original study which the dataset (GSE78956) was generated. The blood sample displays the median methylation of the average methylation of the DMR across 100 control blood samples, while disease sample display mean methylation at the DMR as an individual sample. MO, Methylation Origin; Chr, Chromosome; Est, Established; G, Germline; M, Maternal, P; Paternal; MI, Methylation Index; DMR, Differentially methylated region; TS, Temple Syndrome.

Table A2

Imprinted DMR	Est	MO	Chr	Start	End	Blood	TNDM 1	MI Upper	MI Lower
FAM50B	U	M	chr6	3849082	3850359	58.02	55.57	66.41	48.06
PLAGL1	G	M	chr6	144328078	144329888	63.66	**9.88**	68.48	58.34
PEG10	G	M	chr7	94284759	94287960	43.14	41.18	46.70	39.41
HTR5A	G	M	chr7	154862719	154863382	56.08	57.13	65.35	46.47
GRB10g	G	M	chr7	50848726	50851312	61.75	58.43	70.59	52.79
MEST	G	M	chr7	130130122	130134388	63.44	58.15	69.33	56.40
KvDMR1	G	M	chr11	2719948	2722259	57.98	56.37	66.04	49.04

Table A2 showing an expected epimutation at the *PLAGL1* DMR in a Transient neonatal diabetes mellitus 1 patient samples. Methylation across imprinted DMRs associated with the respective diseases, methylation values with red underlining text showing gains of methylation outside the defined methylation index (MI) while blue underlined text shows loss of methylation outside the defined MI. The patient shows clear hypermethylation at the *PLAGL1* which is the disease-causing epimutation for this disorder. This matches the original study from which this dataset (GSE78956) was derived. The blood sample displays the median methylation of the average methylation of the DMR across 100 control blood samples, while disease sample display mean methylation at the DMR as an individual sample. MO, Methylation Origin; Chr, Chromosome; Est, Established; G, Germ-line; M, Maternal, P; Paternal; MI, Methylation Index; DMR, Differentially methylated region; TNDM1, Transient neonatal diabetes mellitus 1.

Figure A1

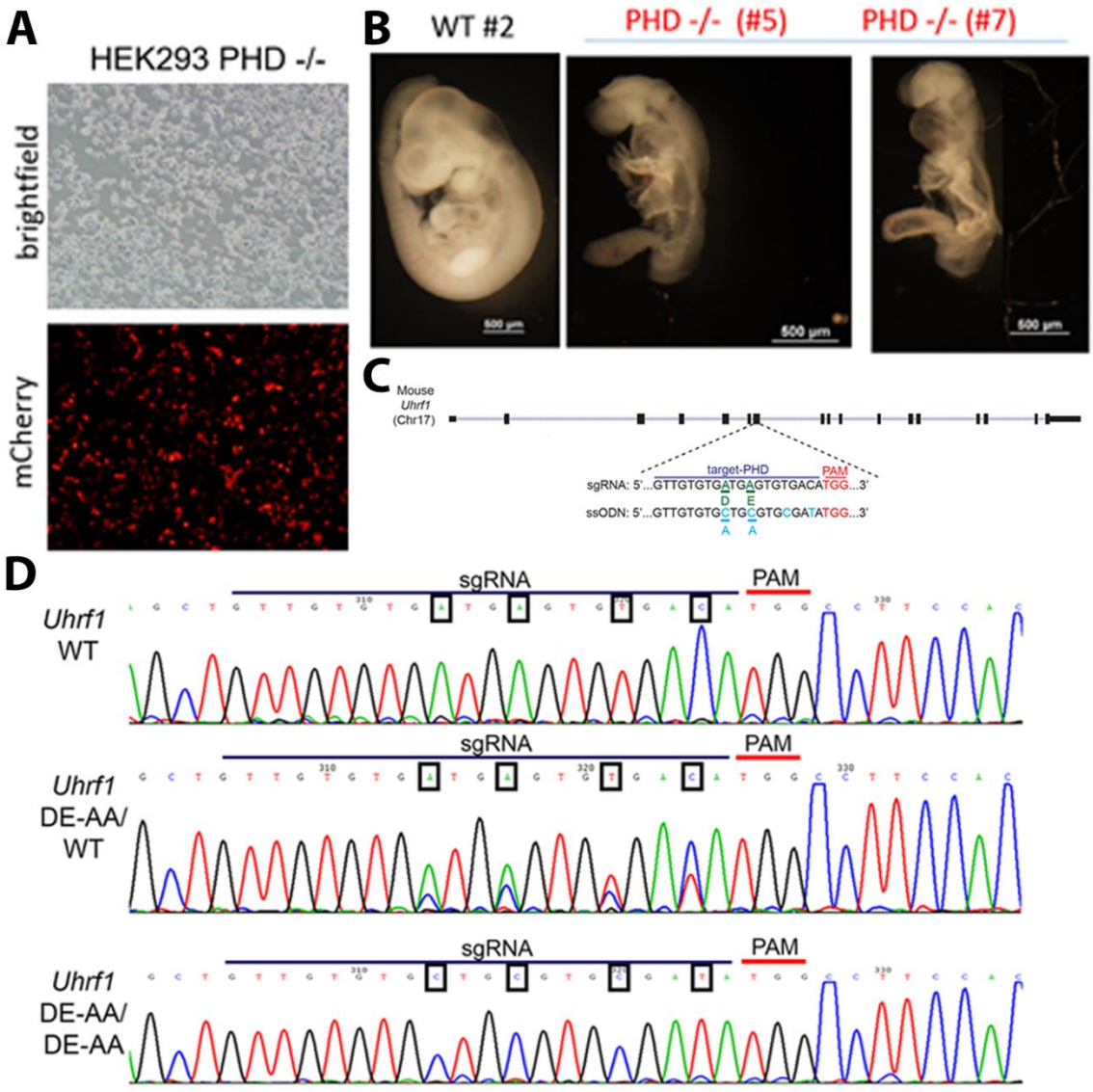


Figure A1: CRISPR models of UHRF1 PHD mutant models in both human embryonic kidney cells (HEK293) and mice. (A) PHD mutation introduced in the HEK293 cells, through CRISPR was successfully carried out resulting in mCHERRY expression. However, single-cell colonies would not grow suggesting lethality. **(B)** Figure as seen in 5.1. **(C)** sgRNA targeting the PHD domain of mouse *Uhrf1*. **(D)** DNA sequence trace from the tail DNA of a living WT, heterozygotic and homozygotic mouse embryos. The boxed sgRNA residues are the induced point mutations caused by the CRISPR.

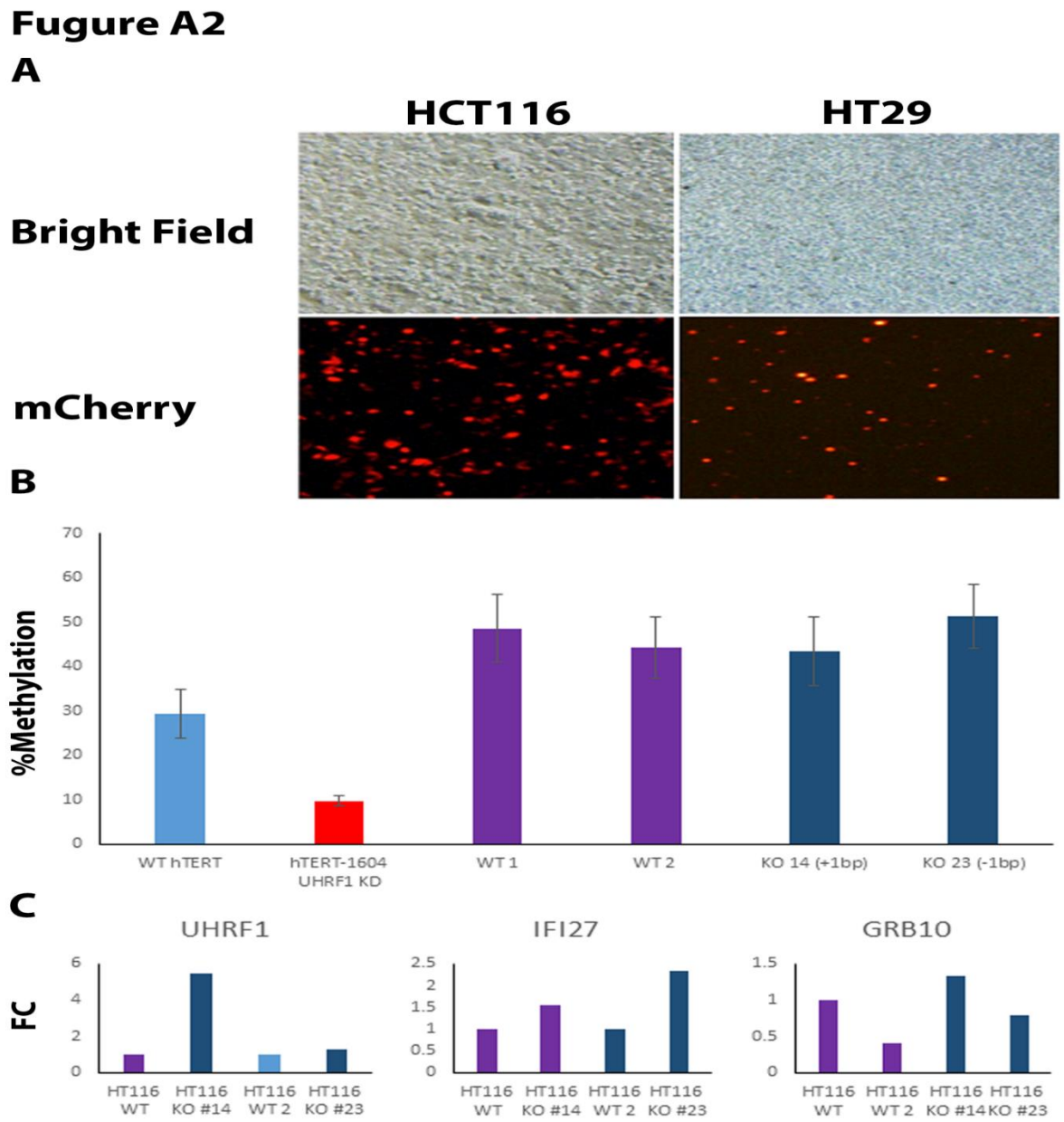


Figure A2: UHRF1 KO clones do not show loss of methylation or changes of expression at methylation-controlled regions. The human colorectal cancer cell lines HCT116 and HT29 were used to try and generate a UHRF1 knockout (KO) differentiated human cell line model through the CRISPR-CAS9 technique. **(A)** After incubation with the plasmid to induce the UHRF1 KO, cells were identified using mCherry and separated using fluorescence-activated cell sorting (FACS). Cells were grown from a single cell, resulting in two HCT116 KOs containing a homozygotic gain (KO 14 (+1bp)) and loss (KO 23 (-1bp)) of a single base pair when the genomic sequence was analysed. **(B)** Using *LINE1* as an indicator of genome-wide methylation there was no sign of loss of methylation for either KO14 and KO23 in comparison to the HCT116 WTs using pyrosequencing. In contrast, UHRF1 depletion resulted in a clear loss of methylation at the *LINE-1* elements in UH4 compared to the hTERT-1604 WT. **(C)** Gene expression analysis using qPCR showed overexpression of UHRF1 when compared to the WT. However, IFI27 and GRB10 which are observed to be controlled by methylation show no alterations to their expression. As both clonal cells did not appear to cause any significant changes of methylation or expression no repeats were carried out and therefore statistical analysis for pyrosequencing assay was carried out using the value of each CpG site for 1 assay. Similarly, the qPCR, only 1 run was completed so the analysis was carried out on the results from the triplicate samples. Significance carried out with a two-tailed T-test with the following key, *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

References

Abreu, A. P., Dauber, A., Macedo, D. B., Noel, S. D., Brito, V. N., Gill, J. C., Cukier, P., Thompson, I. R., Navarro, V. M., Gagliardi, P. C., et al. (2013). Central Precocious Puberty Caused by Mutations in the Imprinted Gene MKRN3. *N. Engl. J. Med.* **368**, 2467–2475.

Acloque, H., Adams, M. S., Fishwick, K., Bronner-Fraser, M. and Nieto, M. A. (2009). Epithelial-mesenchymal transitions: The importance of changing cell state in development and disease. *J. Clin. Invest.* **119**, 1438–1449.

Agostini, L., Martinon, F., Burns, K., Mcdermott, M. F., Hawkins, P. N. and Rg Tschopp, J. (2004). *NALP3 Forms an IL-1-Processing Inflammasome with Increased Activity in Muckle-Wells Autoinflammatory Disorder containing protein called ASC binds and activates pro-caspase-1 (Martinon et al ASC contains a C-terminal CARD motif as well as an N-terminal CARD-like Pyrin domain (Bertin and DiStefano, 2000; Martinon et al., 2001; Masu*.

Algar, E. (2019). Germline Epigenetic Testing of Imprinting Disorders in a Diagnostic Setting. In *Clinical Epigenetics*, pp. 193–215. Springer Singapore.

Allen, N. D., Barton, S. C., Hilton, K., Norris, M. L. and Surani, M. A. (1994). A functional analysis of imprinting in parthenogenetic embryonic stem cells. *Development* **120**, 1473–1482.

Altshuler, D. M., Gibbs, R. A., Peltonen, L., Schaffner, S. F., Yu, F., Dermitzakis, E., Bonnen, P. E., De Bakker, P. I. W., Deloukas, P., Gabriel, S. B., et al. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58.

Altug-Teber, Ö., Dufke, A., Poths, S., Mau-Holzmann, U. A., Bastepe, M., Colleaux, L., Cormier-Daire, V., Eggermann, T., Gillessen-Kaesbach, G., Bonin, M., et al. (2005). A rapid microarray based whole genome analysis for detection of uniparental disomy. *Hum. Mutat.* **26**, 153–159.

Angelman, H. (1965). 'Puppet' Children A Report on Three Cases. *Dev. Med. Child Neurol.* **7**, 681–688.

Anwar, S. L., Krech, T., Hasemeier, B., Schipper, E., Schweitzer, N., Vogel, A., Kreipe, H. and Lehmann, U. (2012). Loss of imprinting and allelic switching at the DLK1-MEG3 locus in human hepatocellular carcinoma. *PLoS One* **7**, e49462.

Arand, J., Spieler, D., Karius, T., Branco, M. R., Meilinger, D., Meissner, A., Jenuwein, T., Xu, G., Leonhardt, H., Wolf, V., et al. (2012). In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLoS Genet.* **8**,.

Arechederra, M., Daian, F., Yim, A., Bazai, S. K., Richelme, S., Dono, R., Saurin, A. J., Habermann, B. H. and Maina, F. (2018). Hypermethylation of gene body CpG islands predicts high dosage of functional oncogenes in liver cancer. *Nat. Commun.* **9**,.

Arima, T. (2005). ZAC, LIT1 (KCNQ1OT1) and p57KIP2 (CDKN1C) are in an imprinted gene network that may play a role in Beckwith-Wiedemann syndrome. *Nucleic Acids Res.* **33**, 2650–2660.

Arita, K., Isogai, S., Oda, T., Unoki, M., Sugita, K., Sekiyama, N., Kuwata, K., Hamamoto, R., Tochio, H., Sato, M., et al. (2012). Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 12950–12955.

Artus, J. and Cohen-Tannoudji, M. (2008). Cell cycle regulation during early mouse embryogenesis. *Mol. Cell. Endocrinol.* **282**, 78–86.

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D. and Irizarry, R. A. (2014). Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369.

Arzenani, M. K., Zade, A. E., Ming, Y., Vijverberg, S. J. H., Zhang, Z., Khan, Z., Sadique, S.,

Kallenbach, L., Hu, L., Vukojevic, V., et al. (2011). Genomic DNA Hypomethylation by Histone Deacetylase Inhibition Implicates DNMT1 Nuclear Dynamics. *Mol. Cell. Biol.* **31**, 4119–4128.

Assenov, Y., Müller, F., Lutsik, P., Walter, J., Lengauer, T. and Bock, C. (2014). Comprehensive analysis of DNA methylation data with RnBeads. *Nat. Methods* **11**, 1138–1140.

Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., et al. (2015). A global reference for human genetic variation. *Nature* **526**, 68–74.

Avvakumov, G. V., Walker, J. R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C. H. and Dhe-Paganon, S. (2008). Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature* **455**, 822–825.

Azzi, S., Rossignol, S., Steunou, V., Sas, T., Thibaud, N., Danton, F., Le Jule, M., Heinrichs, C., Cabrol, S., Gicquel, C., et al. (2009). Multilocus methylation analysis in a large cohort of 11p15-related foetal growth disorders (Russell Silver and Beckwith Wiedemann syndromes) reveals simultaneous loss of methylation at paternal and maternal imprinted loci. *Hum. Mol. Genet.* **18**, 4724–4733.

Azzi, S., Salem, J., Thibaud, N., Chantot-Bastaraud, S., Lieber, E., Netchine, I. and Harbison, M. D. (2015). A prospective study validating a clinical scoring system and demonstrating phenotypical-genotypical correlations in Silver-Russell syndrome. *J. Med. Genet.* **52**, 446–453.

Babbio, F., Pistore, C., Curti, L., Castiglioni, I., Kunderfranco, P., Brino, L., Oudet, P., Seiler, R., Thalman, G. N., Roggero, E., et al. (2012). The SRA protein UHRF1 promotes epigenetic crosstalks and is involved in prostate cancer progression. *Oncogene* **31**, 4878–4887.

Baccarelli, A., Wright, R., Bollati, V., Litonjua, A., Zanobetti, A., Tarantini, L., Sparrow, D.,

Vokonas, P. and Schwartz, J. (2010). Ischemic heart disease and stroke in relation to blood DNA methylation. *Epidemiology* **21**, 819–828.

Baets, J., Duan, X., Wu, Y., Smith, G., Seeley, W. W., Mademan, I., McGrath, N. M., Beadell, N. C., Khoury, J., Botuyan, M. V., et al. (2015). Defects of mutant DNMT1 are linked to a spectrum of neurological disorders. *Brain* **138**, 845–861.

Bak, M., Boonen, S. E., Dahl, C., Hahnemann, J. M. D., Mackay, D. J. D. G., Tümer, Z., Grønskov, K., Temple, I. K., Guldberg, P. and Tommerup, N. (2016). Genome-wide DNA methylation analysis of transient neonatal diabetes type 1 patients with mutations in ZFP57. *BMC Med. Genet.* **17**, 1–8.

Balaton, B. P., Cotton, A. M. and Brown, C. J. (2015). Derivation of consensus inactivation status for X-linked genes from genome-wide studies. *Biol. Sex Differ.* **6**, 35.

Banerji, J., Rusconi, S. and Schaffner, W. (1981). Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299–308.

Bansal, A. and Pinney, S. E. (2017). DNA methylation and its role in the pathogenesis of diabetes. *Pediatr. Diabetes* **18**, 167–177.

Baple, E. L., Poole, R. L., Mansour, S., Willoughby, C., Temple, I. K., Docherty, L. E., Taylor, R. and MacKay, D. J. G. (2011). An atypical case of hypomethylation at multiple imprinted loci. *Eur. J. Hum. Genet.* **19**, 360–362.

Baran, Y., Subramaniam, M., Biton, A., Tukiainen, T., Tsang, E. K., Rivas, M. A., Pirinen, M., Gutierrez-Arcelus, M., Smith, K. S., Kukurba, K. R., et al. (2015). The landscape of genomic imprinting across diverse adult human tissues. *Genome Res.* **25**, 927–936.

Barau, J., Teissandier, A., Zamudio, N., Roy, S., Nalesso, V., Hérault, Y., Guillou, F. and Bourc'his, D. (2016). The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science (80-.).* **354**, 909–912.

Barel, O., Shalev, S. A., Ofir, R., Cohen, A., Zlotogora, J., Shorer, Z., Mazor, G., Finer, G., Khateeb, S., Zilberberg, N., et al. (2008). Maternally Inherited Birk Barel Mental Retardation Dysmorphism Syndrome Caused by a Mutation in the Genomically Imprinted Potassium Channel KCNK9. *Am. J. Hum. Genet.* **83**, 193–199.

Barletta, J. M., Rainier, S. and Feinberg, A. P. (1997). Reversal of loss of imprinting in tumor cells by 5-aza-2'-deoxycytidine. *Cancer Res.* **57**, 48–50.

Barlow, D. P. (1993). Methylation and imprinting: From host defense to gene regulation? *Science (80-.).* **260**, 309–310.

Barlow, D. P. and Bartolomei, M. S. (2014). Genomic imprinting in mammals. *Cold Spring Harb. Perspect. Biol.* **6**, a018382.

Barrangou, R., Birmingham, A., Wiemann, S., Beijersbergen, R. L., Hornung, V. and Van Brabant Smith, A. (2015). Advances in CRISPR-Cas9 genome engineering: lessons learned from RNA interference. *Nucleic Acids Res.* **43**, 3407–3419.

Barrès, R., Osler, M. E., Yan, J., Rune, A., Fritz, T., Caidahl, K., Krook, A. and Zierath, J. R. (2009). Non-CpG Methylation of the PGC-1α Promoter through DNMT3B Controls Mitochondrial Density. *Cell Metab.* **10**, 189–198.

Bartholdi, D., Krajewska-Walasek, M., Õunap, K., Gaspar, H., Chrzanowska, K. H., Ilyana, H., Kayserili, H., Lurie, I. W., Schinzel, A. and Baumer, A. (2009). Epigenetic mutations of the imprinted IGF2-H19 domain in Silver-Russell syndrome (SRS): Results from a large cohort of patients with SRS and SRS-like phenotypes. *J. Med. Genet.* **46**, 192–197.

Bartolomei, M. S. and Ferguson-Smith, A. C. (2011). Mammalian genomic imprinting. *Cold Spring Harb. Perspect. Biol.* **3**, 1–17.

Bartolomei, M. S., Webber, A. L., Brunkow, M. E. and Tilghman, S. M. (1993). Epigenetic mechanisms underlying the imprinting of the mouse H19 gene. *Genes Dev.* **7**, 1663–

1673.

Bashtrykov, P., Jankevicius, G., Jurkowska, R. Z., Ragozin, S. and Jeltsch, A. (2014). The UHRF1 protein stimulates the activity and specificity of the maintenance DNA methyltransferase DNMT1 by an allosteric mechanism. *J. Biol. Chem.* **289**, 4106–4115.

Baylin, S. B. and Jones, P. A. (2016). Epigenetic determinants of cancer. *Cold Spring Harb. Perspect. Biol.* **8**, a019505.

Beach, S. R. H., Dogan, M. V., Lei, M.-K., Cutrona, C. E., Gerrard, M., Gibbons, F. X., Simons, R. L., Brody, G. H. and Philibert, R. A. (2015). Methylomic Aging as a Window onto the Influence of Lifestyle: Tobacco and Alcohol Use Alter the Rate of Biological Aging. *J. Am. Geriatr. Soc.* **63**, 2519–2525.

Beck, C. R., Garcia-Perez, J. L., Badge, R. M. and Moran, J. V. (2011). LINE-1 Elements in Structural Variation and Disease. *Annu. Rev. Genomics Hum. Genet.* **12**, 187–215.

Beckwith, J. (1963). Extreme cytomegaly of the adrenal fetal cortex, hyperplasia of the kidneys and pancreas, and Leydig-cell hyperplasia : another syndrome? *Annu. Meet. West. Soc. Pediatr. Res. Los Angeles, CA, Novemb. 1963*.

Begemann, M., Spengler, S., Gogiel, M., Grasshoff, U., Bonin, M., Betz, R. C., Dufke, A., Spier, I. and Eggermann, T. (2012). Clinical significance of copy number variations in the 11p15.5 imprinting control regions: New cases and review of the literature. *J. Med. Genet.* **49**, 547–553.

Begemann, M., Rezwan, F. I., Beygo, J., Docherty, L. E., Kolarova, J., Schroeder, C., Buiting, K., Chokkalingam, K., Degenhardt, F., Wakeling, E. L., et al. (2018). Maternal variants in NLRP and other maternal effect proteins are associated with multilocus imprinting disturbance in offspring. *J. Med. Genet.* **55**, 497–504.

Bell, R. E., Golan, T., Sheinboim, D., Malcov, H., Amar, D., Salamon, A., Liron, T., Gelfman,

S., Gabet, Y., Shamir, R., et al. (2016). Enhancer methylation dynamics contribute to cancer plasticity and patient mortality. *Genome Res.* **26**, 601–611.

Bens, S., Kolarova, J., Beygo, J., Buiting, K., Caliebe, A., Eggermann, T., Gillessen-Kaesbach, G., Prawitt, D., Thiele-Schmitz, S., Begemann, M., et al. (2016). Phenotypic spectrum and extent of DNA methylation defects associated with multilocus imprinting disturbances. *Epigenomics* **8**, 801–816.

Bens, S., Luedeke, M., Richter, T., Graf, M., Kolarova, J., Barbi, G., Lato, K., Barth, T. F. and Siebert, R. (2017). Mosaic genome-wide maternal isodiploidy: an extreme form of imprinting disorder presenting as prenatal diagnostic challenge. *Clin. Epigenetics* **9**, 111.

Berletch, J. B., Yang, F., Xu, J., Carrel, L. and Disteche, C. M. (2011). Genes that escape from X inactivation. *Hum. Genet.* **130**, 237–245.

Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T. and Schreiber, S. L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 8695–8700.

Bervini, S. and Herzog, H. (2013). Mouse models of Prader–Willi Syndrome: A systematic review. *Front. Neuroendocrinol.* **34**, 107–119.

Beygo, J., Elbracht, M., De Groot, K., Begemann, M., Kanber, D., Platzer, K., Gillessen-Kaesbach, G., Vierzig, A., Green, A., Heller, R., et al. (2015). Novel deletions affecting the MEG3-DMR provide further evidence for a hierarchical regulation of imprinting in 14q32. *Eur. J. Hum. Genet.* **23**, 180–188.

Beygo, J., Küchler, A., Gillessen-Kaesbach, G., Albrecht, B., Eckle, J., Eggermann, T., Gellhaus, A., Kanber, D., Kordaß, U., Lüdecke, H. J., et al. (2017). New insights into the imprinted MEG8-DMR in 14q32 and clinical and molecular description of novel patients with Temple syndrome. *Eur. J. Hum. Genet.* **25**, 935–945.

Beygo, J., Bürger, J., Strom, T. M., Kaya, S. and Buiting, K. (2019a). Disruption of KCNQ1 prevents methylation of the ICR2 and supports the hypothesis that its transcription is necessary for imprint establishment. *Eur. J. Hum. Genet.*

Beygo, J., Buiting, K., Ramsden, S. C., Ellis, R., Clayton-Smith, J. and Kanber, D. (2019b). Update of the EMQN/ACGS best practice guidelines for molecular analysis of Prader-Willi and Angelman syndromes. *Eur. J. Hum. Genet.* **27**, 1326–1340.

Beygo, J., Grosser, C., Kaya, S., Mertel, C., Buiting, K. and Horsthemke, B. (2020). Common genetic variation in the Angelman syndrome imprinting centre affects the imprinting of chromosome 15. *Eur. J. Hum. Genet.* 1–5.

Bhogal, B., Arnaudo, A., Dymkowski, A., Best, A. and Davis, T. L. (2004). Methylation at mouse Cdkn1c is acquired during postimplantation development and functions to maintain imprinted expression. *Genomics* **84**, 961–970.

Bhutani, N., Burns, D. M. and Blau, H. M. (2011). DNA Demethylation Dynamics. *Cell* **146**, 866–872.

Bielinska, B., Blaydes, S. M., Buiting, K., Yang, T., Krajewska-Walasek, M., Horsthemke, B. and Brannan, C. I. (2000). De novo deletions of SNRPN exon 1 in early human and mouse embryos result in a paternal to maternal imprint switch [see comments] [published erratum appears in Nat Genet 2000 Jun;25(2):241]. *Nat Genet* **25**, 74–78.

Bird, A. P. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* **8**, 1499–1504.

Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature* **321**, 209–213.

Bjornsson, H. T., Albert, T. J., Ladd-Acosta, C. M., Green, R. D., Rongione, M. A., Middle, C. M., Irizarry, R. A., Broman, K. W. and Feinberg, A. P. (2008). SNP-specific array-based

allele-specific expression analysis. *Genome Res.* **18**, 771–9.

Blagitko, N., Mergenthaler, S., Schulz, U., Wollmann, H. A., Craigen, W., Eggermann, T., Ropers, H.-H. and Kalscheuer, V. M. (2000). Human GRB10 is imprinted and expressed from the paternal and maternal allele in a highly tissue- and isoform-specific fashion. *Hum. Mol. Genet.* **9**, 1587–1595.

Blanco-Betancourt, C. E., Moncla, A., Milili, M., Jiang, Y. L., Viegas-Péquignot, E. M., Roquelaure, B., Thuret, I. and Schiff, C. (2004). Defective B-cell-negative selection and terminal differentiation in the ICF syndrome. *Blood* **103**, 2683–2690.

Bliek, J., Verde, G., Callaway, J., Maas, S. M., De Crescenzo, A., Sparago, A., Cerrato, F., Russo, S., Ferraiuolo, S., Rinaldi, M. M., et al. (2009). Hypomethylation at multiple maternally methylated imprinted regions including PLAGL1 and GNAS loci in Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.* **17**, 611–619.

Bocklandt, S., Lin, W., Sehl, M. E., Sánchez, F. J., Sinsheimer, J. S., Horvath, S. and Vilain, E. (2011). Epigenetic Predictor of Age. *PLoS One* **6**, e14821.

Bogutz, A. B., Brind'Amour, J., Kobayashi, H., Jensen, K. N., Nakabayashi, K., Imai, H., Lorincz, M. C. and Lefebvre, L. (2019). Evolution of imprinting via lineage-specific insertion of retroviral promoters. *Nat. Commun.* **10**,.

Boonen, S. E., Mackay, D. J. G., Hahnemann, J. M. D., Docherty, L., Gronskov, K., Lehmann, A., Larsen, L. G., Haemers, A. P., Kockaerts, Y., Dooms, L., et al. (2013). Transient neonatal diabetes, ZFP57, and hypomethylation of multiple imprinted loci. *Diabetes Care* **36**, 505–512.

Borchiellini, M., Ummarino, S. and Di Ruscio, A. (2019). The Bright and Dark Side of DNA Methylation: A Matter of Balance. *Cells* **8**, 1243.

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., Forné, T. and Weber, M.

(2010). Targets and dynamics of promoter DNA methylation during early mouse development. *Nat. Genet.* **42**, 1093–1100.

Bostick, M., Jong, K. K., Estève, P. O., Clark, A., Pradhan, S. and Jacobsen, S. E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science (80-.).* **317**, 1760–1764.

Bourc'his, D. and Bestor, T. H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431**, 96–99.

Bourc'his, D., Xu, G. L., Lin, C. S., Bollman, B. and Bestor, T. H. (2001). Dnmt3L and the Establishment of Maternal Genomic Imprints. *Science (80-.).* **294**, 2536–2539.

Brioude, F., Oliver-Petit, I., Blaise, A., Praz, F., Rossignol, S., Le Jule, M., Thibaud, N., Faussat, A. M., Tauber, M., Le Bouc, Y., et al. (2013). CDKN1C mutation affecting the PCNA-binding domain as a cause of familial Russell Silver syndrome. *J. Med. Genet.* **50**, 823–830.

Brioude, F., Kalish, J. M., Mussa, A., Foster, A. C., Bliek, J., Ferrero, G. B., Boonen, S. E., Cole, T., Baker, R., Bertoletti, M., et al. (2018). Clinical and molecular diagnosis, screening and management of Beckwith-Wiedemann syndrome: An international consensus statement. *Nat. Rev. Endocrinol.* **14**, 229–249.

Bronner, C., Krifa, M. and Mousli, M. (2013). Increasing role of UHRF1 in the reading and inheritance of the epigenetic code as well as in tumorogenesis. *Biochem. Pharmacol.* **86**, 1643–1649.

Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R. and Willard, H. F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**, 38–44.

Bruey, J. M., Bruey-Sedano, N., Newman, R., Chandler, S., Stehlik, C. and Reed, J. C. (2004).

PAN1/NALP2/PYPAF2, an Inducible Inflammatory Mediator That Regulates NF-κB and Caspase-1 Activation in Macrophages. *J. Biol. Chem.* **279**, 51897–51907.

Bruno, S., Bochicchio, M. T., Franchini, E., Padella, A., Marconi, G., Ghelli Luserna Di Rorà, A., Venturi, C., Raffini, M., Prisinzano, G., Ferrari, A., et al. (2019). Identification of Two DNMT3A Mutations Compromising Protein Stability and Methylation Capacity in Acute Myeloid Leukemia. *J. Oncol.* **2019**,.

Buiting, K. (2010). Prader-Willi syndrome and Angelman syndrome. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **154**, 365–376.

Buiting, K., Williams, C. and Horsthemke, B. (2016). Angelman syndrome-insights into a rare neurogenetic disorder. *Nat. Rev. Neurol.* **12**, 584–593.

Butler, M. G., Bittel, D. C., Kibiryeva, N., Talebizadeh, Z. and Thompson, T. (2004). Behavioral Differences among Subjects with Prader-Willi Syndrome and Type I or Type II Deletion and Maternal Disomy. *Pediatrics* **113**, 565–573.

Butler, M. G., Manzardo, A. M. and Forster, J. L. (2016). Prader-Willi Syndrome: Clinical Genetics and Diagnostic Aspects with Treatment Approaches. *Curr. Pediatr. Rev.* **12**, 136–66.

Caffrey, A., Irwin, R. E., McNulty, H., Strain, J. J., Lees-Murdock, D. J., McNulty, B. A., Ward, M., Walsh, C. P. and Pentieva, K. (2018). Gene-specific DNA methylation in newborns in response to folic acid supplementation during the second and third trimesters of pregnancy: epigenetic analysis from a randomized controlled trial. *Am. J. Clin. Nutr.* **107**, 566–575.

Cai, Y. N., Dai, X. H., Zhang, Q. H., Hu, R. and Dai, Z. M. (2015). Gene expression profiling of somatic and pluripotent cells reveals novel pathways involved in reprogramming. *Genet. Mol. Res.* **14**, 12085–12092.

Cai, Y., Tsai, H. C., Yen, R. W. C., Zhang, Y. W., Kong, X., Wang, W., Xia, L. and Baylin, S. B. (2017). Critical threshold levels of DNA methyltransferase 1 are required to maintain DNA methylation across the genome in human cancer cells. *Genome Res.* **27**, 533–544.

Callebaut, I., Courvalin, J. C. and Mornon, J. P. (1999). The BAH (bromo-adjacent homology) domain: a link between DNA methylation, replication and transcriptional regulation. *FEBS Lett.* **446**, 189–93.

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F. J., et al. (2019a). The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502.

Cao, Y., Li, M., Liu, F., Ni, X., Wang, S., Zhang, H., Sui, X. and Huo, R. (2019b). Deletion of maternal UHRF1 severely reduces mouse oocyte quality and causes developmental defects in preimplantation embryos. *FASEB J.* **33**, 8294–8305.

Carmona, J. J., Accomando, W. P., Binder, A. M., Hutchinson, J. N., Pantano, L., Izzi, B., Just, A. C., Lin, X., Schwartz, J., Vokonas, P. S., et al. (2017). Empirical comparison of reduced representation bisulfite sequencing and Infinium BeadChip reproducibility and coverage of DNA methylation in humans. *npj Genomic Med.* **2**, 13.

Catchpoole, D., Lam, W. W. K., Valler, D., Temple, I. K., Joyce, J. A., Reik, W., Schofield, P. N. and Maher, E. R. (1997). Epigenetic modification and uniparental inheritance of H19 in Beckwith-Wiedemann syndrome. *J. Med. Genet.* **34**, 353–359.

Chaitankar, V., Karakülah, G., Ratnapriya, R., Giuste, F. O., Brooks, M. J. and Swaroop, A. (2016). Next generation sequencing technology and genomewide data analysis: Perspectives for retinal research. *Prog. Retin. Eye Res.* **55**, 1–31.

Chamberlain, S. J. and Lalande, M. (2010a). Neurodevelopmental disorders involving genomic imprinting at human chromosome 15q11-q13. *Neurobiol. Dis.* **39**, 13–20.

Chamberlain, S. J. and Lalande, M. (2010b). Angelman syndrome, a genomic imprinting disorder of the brain. *J. Neurosci.* **30**, 9958–9963.

Chang, S. and Bartolomei, M. S. (2020). Modeling human epigenetic disorders in mice: Beckwith-Wiedemann syndrome and Silver-Russell syndrome. *DMM Dis. Model. Mech.* **13**,.

Charalambous, M., Smith, F. M., Bennett, W. R., Crew, T. E., Mackenzie, F. and Ward, A. (2003). Disruption of the imprinted Grb10 gene leads to disproportionate overgrowth by an Igf2-independent mechanism. *Proc. Natl. Acad. Sci.* **100**, 8292–8297.

Chen, B. F. and Chan, W. Y. (2014). The de novo DNA methyltransferase DNMT3A in development and cancer. *Epigenetics* **9**, 669–677.

Chen, T., Ueda, Y., Xie, S. and Li, E. (2002). A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *J. Biol. Chem.* **277**, 38746–38754.

Chen, T., Ueda, Y., Dodge, J. E., Wang, Z. and Li, E. (2003). Establishment and Maintenance of Genomic Methylation Patterns in Mouse Embryonic Stem Cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* **23**, 5594.

Chen, L., Chen, K., Lavery, L. A., Baker, S. A., Shaw, C. A., Li, W. and Zoghbi, H. Y. (2015). MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 5509–5514.

Chen, C. L., Lee, C. N., Lin, M. W., Hsu, W. W., Tai, Y. Y. and Lin, S. Y. (2019). Prenatal diagnosis of paternal uniparental disomy for chromosome 14 using a single-nucleotide-polymorphism-based microarray analysis: A case report. *J. Formos. Med. Assoc.* **118**, 739–742.

Cheng, J., Yang, Y., Fang, J., Xiao, J., Zhu, T., Chen, F., Wang, P., Li, Z., Yang, H. and Xu, Y.

(2013). Structural insight into coordinated recognition of trimethylated histone H3 lysine 9 (H3K9me3) by the plant homeodomain (PHD) and tandem tudor domain (TTD) of UHRF1 (ubiquitin-like, containing PHD and RING finger domains, 1) protein. *J. Biol. Chem.* **288**, 1329–1339.

Cheong, C. Y., Chng, K., Ng, S., Chew, S. B., Chan, L. and Ferguson-Smith, A. C. (2015). Germline and somatic imprinting in the nonhuman primate highlights species differences in oocyte methylation. *Genome Res.* **125**, 611–623.

Chernyavskaya, Y., Mudbhary, R., Zhang, C., Tokarz, D., Jacob, V., Gopinath, S., Sun, X., Wang, S., Magnani, E., Madakashira, B. P., et al. (2017). Loss of dna methylation in zebrafish embryos activates retrotransposons to trigger antiviral signaling. *Dev.* **144**, 2925–2939.

Chouliaras, L., Mastroeni, D., Delvaux, E., Grover, A., Kenis, G., Hof, P. R., Steinbusch, H. W. M., Coleman, P. D., Rutten, B. P. F. and van den Hove, D. L. A. (2013). Consistent decrease in global DNA methylation and hydroxymethylation in the hippocampus of Alzheimer's disease patients. *Neurobiol. Aging* **34**, 2091–2099.

Citterio, E., Papait, R., Nicassio, F., Vecchi, M., Gomiero, P., Mantovani, R., Di Fiore, P. P. and Bonapace, I. M. (2004). Np95 Is a Histone-Binding Protein Endowed with Ubiquitin Ligase Activity. *Mol. Cell. Biol.* **24**, 2526–2535.

Clayton-Smith, J. and Laan, L. (2003). Angelman syndrome: A review of the clinical and genetic aspects. *J. Med. Genet.* **40**, 87–95.

Clement-Jones, M. (2000). The short stature homeobox gene SHOX is involved in skeletal abnormalities in Turner syndrome. *Hum. Mol. Genet.* **9**, 695–702.

Cohen, C. J., Lock, W. M. and Mager, D. L. (2009). Endogenous retroviral LTRs as promoters for human genes: A critical assessment. *Gene* **448**, 105–114.

Cooper, S., Dienstbier, M., Hassan, R., Schermelleh, L., Sharif, J., Blackledge, N. P., DeMarco, V., Elderkin, S., Koseki, H., Klose, R., et al. (2014). Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. *Cell Rep.* **7**, 1456–1470.

Copp, A. J. (1979). *Interaction between inner cell mass and trophectoderm of the mouse blastocyst II. The fate of the polar trophectoderm*.

Costello, J. F. and Plass, C. (2001). Methylation matters. *J. Med. Genet.* **38**, 285–303.

Coucouvanis, E. and Martin, G. R. (1995). Signals for death and survival: A two-step mechanism for cavitation in the vertebrate embryo. *Cell* **83**, 279–287.

Coulondre, C., Miller, J. H., Farabaugh, P. J. and Gilbert, W. (1978). Molecular basis of base substitution hotspots in Escherichia coli. *Nature* **274**, 775–780.

Court, F., Martin-Trujillo, A., Romanelli, V., Garin, I., Iglesias-Platas, I., Salafsky, I., Guitart, M., Perez de Nanclares, G., Lapunzina, P. and Monk, D. (2013). Genome-Wide Allelic Methylation Analysis Reveals Disease-Specific Susceptibility to Multiple Methylation Defects in Imprinting Syndromes. *Hum. Mutat.* **34**, 595–602.

Court, F., Tayama, C., Romanelli, V., Martin-Trujillo, A., Iglesias-Platas, I., Okamura, K., Sugahara, N., Simón, C., Moore, H., Harness, J. V, et al. (2014). Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. *Genome Res.* **24**, 554–69.

Cowley, M. and Oakey, R. J. (2010). Retrotransposition and genomic imprinting. *Brief. Funct. Genomics* **9**, 340–6.

Crouse, H. V (1960). The Controlling Element in Sex Chromosome Behavior in Sciara. *Genetics* **45**, 1429–43.

Dauber, A., Cunha-Silva, M., MacEdo, D. B., Brito, V. N., Abreu, A. P., Roberts, S. A., Montenegro, L. R., Andrew, M., Kirby, A., Weirauch, M. T., et al. (2017). Paternally Inherited DLK1 deletion associated with familial central precocious puberty. *J. Clin. Endocrinol. Metab.* **102**, 1557–1567.

Dawlaty, M. M., Ganz, K., Powell, B. E., Hu, Y. C., Markoulaki, S., Cheng, A. W., Gao, Q., Kim, J., Choi, S. W., Page, D. C., et al. (2011). Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9**, 166–175.

Dawlaty, M. M., Breiling, A., Le, T., Raddatz, G., Barrasa, M. I., Cheng, A. W., Gao, Q., Powell, B. E., Li, Z., Xu, M., et al. (2013). Combined Deficiency of Tet1 and Tet2 Causes Epigenetic Abnormalities but Is Compatible with Postnatal Development. *Dev. Cell* **24**, 310–323.

Dawlaty, M. M., Breiling, A., Le, T., Barrasa, M. I., Raddatz, G., Gao, Q., Powell, B. E., Cheng, A. W., Faull, K. F., Lyko, F., et al. (2014). Loss of tet enzymes compromises proper differentiation of embryonic stem cells. *Dev. Cell* **29**, 102–111.

De Crescenzo, A., Sparago, A., Cerrato, F., Palumbo, O., Carella, M., Miceli, M., Bronshtein, M., Riccio, A. and Yaron, Y. (2013). Paternal deletion of the 11p15.5 centromeric-imprinting control region is associated with alteration of imprinted gene expression and recurrent severe intrauterine growth restriction. *J. Med. Genet.* **50**, 99–103.

de Koning, A. P. J., Gu, W., Castoe, T. A., Batzer, M. A. and Pollock, D. D. (2011). Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. *PLoS Genet.* **7**, e1002384.

Deangelis, A. M., Martini, A. E. and Owen, C. M. (2018). Assisted Reproductive Technology and Epigenetics. *Semin. Reprod. Med.* **36**, 221–232.

DeBerardinis, R. J., Goodier, J. L., Ostertag, E. M. and Kazazian, H. H. (1998). Rapid amplification of a retrotransposon subfamily is evolving the mouse genome. *Nat. Genet.*

20, 288–290.

Dedeurwaerder, S., Defrance, M., Calonne, E., Denis, H., Sotiriou, C. and Fuks, F. (2011). Evaluation of the Infinium Methylation 450K technology. *Epigenomics* **3**, 771–784.

Demond, H., Anvar, Z., Jahromi, B. N., Sparago, A., Verma, A., Davari, M., Calzari, L., Russo, S., Jahromi, M. A., Monk, D., et al. (2019). A KHDC3L mutation resulting in recurrent hydatidiform mole causes genome-wide DNA methylation loss in oocytes and persistent imprinting defects post-fertilisation. *Genome Med.* **11**, 84.

Devriendt, K. (2005). Hydatidiform mole and triploidy: the role of genomic imprinting in placental development. *Hum. Reprod. Update* **11**, 137–142.

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R. Z., Ragozin, S. and Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J. Biol. Chem.* **285**, 26114–26120.

Dietzsch, J., Feineis, D. and Höbartner, C. (2018). Chemoselective labeling and site-specific mapping of 5-formylcytosine as a cellular nucleic acid modification. *FEBS Lett.* **592**, 2032–2047.

Ding, C., Huang, S., Qi, Q., Fu, R., Zhu, W., Cai, B., Hong, P., Liu, Z., Gu, T., Zeng, Y., et al. (2015). Derivation of a Homozygous Human Androgenetic Embryonic Stem Cell Line. *Stem Cells Dev.* **24**, 2307–2316.

Docherty, L. E., Rezwan, F. I., Poole, R. L., Jagoe, H., Lake, H., Lockett, G. A., Arshad, H., Wilson, D. I., Holloway, J. W., Temple, I. K., et al. (2014). Genome-wide DNA methylation analysis of patients with imprinting disorders identifies differentially methylated regions associated with novel candidate imprinted genes. *J. Med. Genet.* **51**, 229–238.

Docherty, L. E., Rezwan, F. I., Poole, R. L., Turner, C. L. S., Kivuva, E., Maher, E. R., Smithson,

S. F., Hamilton-Shield, J. P., Patalan, M., Gizewska, M., et al. (2015). Mutations in NLRP5 are associated with reproductive wastage and multilocus imprinting disorders in humans. *Nat. Commun.* **6**,.

Dong, K. B., Maksakova, I. A., Mohn, F., Leung, D., Appanah, R., Lee, S., Yang, H. W., Lam, L. L., Mager, D. L., Schübeler, D., et al. (2008). DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J.* **27**, 2691–2701.

Dong, E., Gavin, D. P., Chen, Y. and Davis, J. (2012). Upregulation of TET1 and downregulation of APOBEC3A and APOBEC3C in the parietal cortex of psychotic patients. *Transl. Psychiatry* **2**,.

Dong, J., Wang, X., Cao, C., Wen, Y., Sakashita, A., Chen, S., Zhang, J., Zhang, Y., Zhou, L., Luo, M., et al. (2019). UHRF1 suppresses retrotransposons and cooperates with PRMT5 and PIWI proteins in male germ cells. *Nat. Commun.* **10**, 1–14.

Duymich, C. E., Charlet, J., Yang, X., Jones, P. A. and Liang, G. (2016). DNMT3B isoforms without catalytic activity stimulate gene body methylation as accessory proteins in somatic cells. *Nat. Commun.* **7**, 11453.

Ecco, G., Cassano, M., Kauzlaric, A., Duc, J., Coluccio, A., Offner, S., Imbeault, M., Rowe, H. M., Turelli, P. and Trono, D. (2016). Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. *Dev. Cell* **36**, 611–623.

Edwards, C. A. and Ferguson-Smith, A. C. (2007). Mechanisms regulating imprinted genes in clusters. *Curr. Opin. Cell Biol.* **19**, 281–289.

Edwards, J. R., Yarychkivska, O., Boulard, M. and Bestor, T. H. (2017). DNA methylation and DNA methyltransferases. *Epigenetics Chromatin* **10**, 23.

Eggermann, T., Heilsberg, A. K., Bens, S., Siebert, R., Beygo, J., Buiting, K., Begemann, M. and Soellner, L. (2014). Additional molecular findings in 11p15-associated imprinting

disorders: An urgent need for multi-locus testing. *J. Mol. Med.* **92**, 769–777.

Eggermann, T., Soellner, L., Buiting, K. and Kotzot, D. (2015a). Mosaicism and uniparental disomy in prenatal diagnosis. *Trends Mol. Med.* **21**, 77–87.

Eggermann, T., Perez de Nanclares, G., Maher, E. R., Temple, I. K., Tümer, Z., Monk, D., Mackay, D. J. G., Grønskov, K., Riccio, A., Linglart, A., et al. (2015b). Imprinting disorders: a group of congenital disorders with overlapping patterns of molecular changes affecting imprinted loci. *Clin. Epigenetics* **7**,.

Eggermann, T., Netchine, I., Temple, I. K., Tümer, Z., Monk, D., Mackay, D., Grønskov, K., Riccio, A., Linglart, A. and Maher, E. R. (2015c). Congenital imprinting disorders: EUCID.net - a network to decipher their aetiology and to improve the diagnostic and clinical care. *Clin. Epigenetics* **7**, 23.

Eggermann, K., Bliek, J., Brioude, F., Algar, E., Buiting, K., Russo, S., Tümer, Z., Monk, D., Moore, G., Antoniadi, T., et al. (2016). EMQN best practice guidelines for the molecular genetic testing and reporting of chromosome 11p15 imprinting disorders: Silver-Russell and Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.* **24**, 1377–1387.

Ehrhart, F., Janssen, K. J. M., Coort, S. L., Evelo, C. T. and Curfs, L. M. G. (2018). Prader-Willi syndrome and Angelman syndrome: Visualisation of the molecular pathways for two chromosomal disorders. *World J. Biol. Psychiatry* 1–13.

Ehrlich, M., Jackson, K. and Weemaes, C. (2006). Immunodeficiency, centromeric region instability, facial anomalies syndrome (ICF). *Orphanet J. Rare Dis.* **1**,.

Elalaoui, S. C., Garin, I., Sefiani, A. and Perez De Nanclares, G. (2014). Maternal hypomethylation of KvDMR in a monozygotic male twin pair discordant for beckwith-wiedemann syndrome. *Mol. Syndromol.* **5**, 41–46.

Elli, F. M., Bordogna, P., Arosio, M., Spada, A. and Mantovani, G. (2018). Mosaicism for GNAS

methylation defects associated with pseudohypoparathyroidism type 1B arose in early post-zygotic phases. *Clin. Epigenetics* **10**, 16.

Engel, E. (1980). A new genetic concept: Uniparental disomy and its potential effect, isodisomy. *Am. J. Med. Genet.* **6**, 137–143.

Engel, E. (1993). Uniparental disomy revisited: The first twelve years. *Am. J. Med. Genet.* **46**, 670–674.

Engel, N., Thorvaldsen, J. L. and Bartolomei, M. S. (2006). CTCF binding sites promote transcription initiation and prevent DNA methylation on the maternal allele at the imprinted H19/Igf2 locus. *Hum. Mol. Genet.* **15**, 2945–2954.

Esteban, M. A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., et al. (2010). Vitamin C Enhances the Generation of Mouse and Human Induced Pluripotent Stem Cells. *Cell Stem Cell* **6**, 71–79.

Etoh, T., Kanai, Y., Ushijima, S., Nakagawa, T., Nakanishi, Y., Sasako, M., Kitano, S. and Hirohashi, S. (2004). Increased DNA methyltransferase 1 (DNMT1) protein expression correlates significantly with poorer tumor differentiation and frequent DNA hypermethylation of multiple CpG islands in gastric cancers. *Am. J. Pathol.* **164**, 689–99.

Farhadova, S., Gomez-Velazquez, M. and Feil, R. (2019). Stability and lability of parental methylation imprints in development and disease. *Genes (Basel).* **10**,.

Feng, Q., Moran, J. V., Kazazian, H. H. and Boeke, J. D. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916.

Feng, G., Hobbs, J., Lu, X., Yu, Y., Du, P., Kibbe, W. A., Chandler, J., Hou, L. and Lin, S. M. (2013). A statistical method to estimate DNA copy number from Illumina high-density methylation arrays. *Syst. Biomed.* **1**, 94–98.

Ferguson-Smith, A. C. (2011). Genomic imprinting: the emergence of an epigenetic paradigm. *Nat. Rev. Genet.* **12**, 565–575.

Ferguson-Smith, A. C., Sasaki, H., Cattanach, B. M. and Surani, M. A. (1993). Parental-origin-specific epigenetic modification of the mouse H19 gene. *Nature* **362**, 751–755.

Feschotte, C. and Pritham, E. J. (2007). DNA Transposons and the Evolution of Eukaryotic Genomes. *Annu. Rev. Genet.* **41**, 331–368.

Finnegan, D. J. (2012). Retrotransposons. *Curr. Biol.* **22**, R432–R437.

Fokstuen, S. and Kotzot, D. (2014). Chromosomal rearrangements in patients with clinical features of Silver-Russell syndrome. *Am. J. Med. Genet. Part A* **164**, 1595–1605.

Fontana, L., Bedeschi, M. F., Maitz, S., Cereda, A., Faré, C., Motta, S., Seresini, A., D'Ursi, P., Orro, A., Pecile, V., et al. (2018). Characterization of multi-locus imprinting disturbances and underlying genetic defects in patients with chromosome 11p15.5 related imprinting disorders. *Epigenetics* **13**, 897–909.

Fradin, D., Cheslack-Postava, K., Ladd-Acosta, C., Newschaffer, C., Chakravarti, A., Arking, D. E., Feinberg, A. and Fallin, M. D. (2010). Parent-of-origin effects in autism identified through genome-wide linkage analysis of 16,000 SNPs. *PLoS One* **5**,.

Fransquet, P. D., Wrigglesworth, J., Woods, R. L., Ernst, M. E. and Ryan, J. (2019). The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis. *Clin. Epigenetics* **11**, 62.

Galerneau, F. (2018). 109-Beckwith-Wiedemann Syndrome. In *Obstetric Imaging: Fetal Diagnosis and Care: Second Edition*, pp. 462-466.e1. Elsevier Inc.

Gao, X., Zhang, Y., Breitling, L. P. and Brenner, H. (2016). Relationship of tobacco smoking and smoking-related DNA methylation with epigenetic age acceleration. *Oncotarget* **7**,

46878–46889.

Gardiner-Garden, M. and Frommer, M. (1987). CpG Islands in vertebrate genomes. *J. Mol. Biol.* **196**, 261–282.

Garfield, A. S., Cowley, M., Smith, F. M., Moorwood, K., Stewart-Cox, J. E., Gilroy, K., Baker, S., Xia, J., Dalley, J. W., Hurst, L. D., et al. (2011). Distinct physiological and behavioural functions for parental alleles of imprinted Grb10. *Nature* **469**, 534–540.

Garrick, D., Fiering, S., Martin, D. I. K. and Whitelaw, E. (1998). Repeat-induced gene silencing in mammals. *Nat. Genet.* **18**, 56–59.

Gelato, K. A., Tauber, M., Ong, M. S., Winter, S., Hiragami-Hamada, K., Sindlinger, J., Lemak, A., Bultsma, Y., Houliston, S., Schwarzer, D., et al. (2014). Accessibility of different histone H3-binding domains of UHRF1 is allosterically regulated by phosphatidylinositol 5-phosphate. *Mol. Cell* **54**, 905–919.

Gendrel, A. V., Apedaile, A., Coker, H., Termanis, A., Zvetkova, I., Godwin, J., Tang, Y. A., Huntley, D., Montana, G., Taylor, S., et al. (2012). Smchd1-Dependent and -Independent Pathways Determine Developmental Dynamics of CpG Island Methylation on the Inactive X Chromosome. *Dev. Cell* **23**, 265–279.

Gibney, E. R. and Nolan, C. M. (2010). Epigenetics and gene expression. *Heredity (Edinb).* **105**, 4–13.

Gicquel, C., Rossignol, S., Cabrol, S., Houang, M., Steunou, V., Barbu, V., Danton, F., Thibaud, N., Le Merrer, M., Burglen, L., et al. (2005). Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. *Nat. Genet.* **37**, 1003–1007.

Glenn, C. C., Porter, K. A., Jong, M. T. C., Nicholls, R. D. and Driscoll, D. J. (1993). Functional imprinting and epigenetic modification of the human SNRPN gene. *Hum. Mol. Genet.* **2**,

2001–2005.

Glenn, C. C., Saitoh, S., Jong, M. T. C., Filbrandt, M. M., Surti, U., Driscoll, D. J., Nicholls3'4 ', R. D. and Nicholls, R. D. (1996). *Gene Structure, DNA Methylation, and Imprinted Expression of the Human SNRPN Gene*.

Gomes, K. M. S., Costa, I. C., Dos Santos, J. F., Dourado, P. M. M., Forni, M. F. and Ferreira, J. C. B. (2017). Induced pluripotent stem cells reprogramming: Epigenetics and applications in the regenerative medicine. *Rev. Assoc. Med. Bras.* **63**, 180–189.

Gordon, C. A., Hartono, S. R. and Chédin, F. (2013). Inactive DNMT3B Splice Variants Modulate De Novo DNA Methylation. *PLoS One* **8**,.

Grafodatskaya, D., Choufani, S., Basran, R. and Weksberg, R. (2016). An Update on Molecular Diagnostic Testing of Human Imprinting Disorders. *J. Pediatr. Genet.* **06**, 003–017.

Grant, M., Zuccotti, M. and Monk, M. (1992). Methylation of CpG sites of two X–linked genes coincides with X–inactivation in the female mouse embryo but not in the germ line. *Nat. Genet.* **2**, 161–166.

Gray, T. A., Saitoh, S. and Nicholls, R. D. (1999). An imprinted, mammalian bicistronic transcript encodes two independent proteins. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 5616–21.

Greißel, A., Culmes, M., Napieralski, R., Wagner, E., Gebhard, H., Schmitt, M., Zimmermann, A., Eckstein, H. H., Zernecke, A. and Pelisek, J. (2015). Alternation of histone and DNA methylation in human atherosclerotic carotid plaques. *Thromb. Haemost.* **114**, 390–402.

Grothaus, K., Kanber, D., Gellhaus, A., Mikat, B., Kolarova, J., Siebert, R., Wieczorek, D. and Horsthemke, B. (2016). Genome-wide methylation analysis of retrocopy-associated CpG islands and their genomic environment. *Epigenetics* **11**, 216–26.

Gu, T. P., Guo, F., Yang, H., Wu, H. P., Xu, G. F., Liu, W., Xie, Z. G., Shi, L., He, X., Jin, S. G., et

al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–612.

Guarrera, S., Fiorito, G., Onland-Moret, N. C., Russo, A., Agnoli, C., Allione, A., Di Gaetano, C., Mattiello, A., Ricceri, F., Chiodini, P., et al. (2015). Gene-specific DNA methylation profiles and LINE-1 hypomethylation are associated with myocardial infarction risk. *Clin. Epigenetics* **7**,.

Guibert, S., Forné, T. and Weber, M. (2012). Global profiling of DNA methylation erasure in mouse primordial germ cells. *Genome Res.* **22**, 633–641.

Gujar, H., Weisenberger, D. J. and Liang, G. (2019). The roles of human DNA methyltransferases and their isoforms in shaping the epigenome. *Genes (Basel).* **10**,.

Gunay-Aygun, M., Schwartz, S., Heeger, S., O'Riordan, M. A. and Cassidy, S. B. (2001). The changing purpose of Prader-Willi syndrome clinical diagnostic criteria and proposed revised criteria. *Pediatrics* **108**,.

Habib, W. A., Brioude, F., Azzi, S., Rossignol, S., Linglart, A., Sobrier, M. L., Giabicani, É., Steunou, V., Harbison, M. D., Le Bouc, Y., et al. (2019). Transcriptional profiling at the DLK1/MEG3 domain explains clinical overlap between imprinting disorders. *Sci. Adv.* **5**,.

Hackett, J. A., Reddington, J. P., Nestor, C. E., Dunican, D. S., Branco, M. R., Reichmann, J., Reik, W., Surani, M. A., Adams, I. R. and Meehan, R. R. (2012). Promoter DNA methylation couples genome-defence mechanisms to epigenetic reprogramming in the mouse germline. *Development* **139**, 3623–32.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J. and Surani, M. A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117**, 15–23.

Hamatani, T., Carter, M. G., Sharov, A. A. and Ko, M. S. H. (2004). Dynamics of global gene

expression changes during mouse preimplantation development. *Dev. Cell* **6**, 117–131.

Hamatani, T., Ko, M. S., Yamada, M., Kuji, N., Mizusawa, Y., Shoji, M., Hada, T., Asada, H., Maruyama, T. and Yoshimura, Y. (2006). Global gene expression profiling of preimplantation embryos. *Hum. cell Off. J. Hum. Cell Res. Soc.* **19**, 98–117.

Hanahan, D. and Robert A, W. (2017). Biological Hallmarks of cancer. *Holland-Frei cancer Med.* **01**, 1–10.

Hanna, C. W. (2020). Placental imprinting: Emerging mechanisms and functions. *PLoS Genet.* **16**,.

Hanna, C. W. and Kelsey, G. (2014). The specification of imprints in mammals. *Heredity (Edinb).* **113**, 176–183.

Hanna, C. W., Pérez-Palacios, R., Gahurova, L., Schubert, M., Krueger, F., Biggins, L., Andrews, S., Colomé-Tatché, M., Bourc'His, D., Dean, W., et al. (2019). Endogenous retroviral insertions drive non-canonical imprinting in extra-embryonic tissues. *Genome Biol.* **20**,.

Hannon, E., Knox, O., Sugden, K., Burrage, J., Wong, C. C. Y., Belsky, D. W., Corcoran, D. L., Arseneault, L., Moffitt, T. E., Caspi, A., et al. (2018). Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. *PLoS Genet.* **14**, e1007544.

Hannula-Jouppi, K., Muurinen, M., Lipsanen-Nyman, M., Reinius, L. E., Ezer, S., Greco, D. and Kere, J. (2013). Differentially methylated regions in maternal and paternal uniparental disomy for chromosome 7. *Epigenetics* **9**, 351–365.

Harrison, J. S., Cornett, E. M., Goldfarb, D., Darosa, P. A., Li, Z. M., Yan, F., Dickson, B. M., Guo, A. H., Cantu, D. V., Kaustov, L., et al. (2016). Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1. *Elife*

5,.

Hatada, I., Ohashi, H., Fukushima, Y., Kaneko, Y., Inoue, M., Komoto, Y., Okada, A., Ohishi, S., Nabetani, A., Morisaki, H., et al. (1996). An imprinted gene p57(KIP2) is mutated in Beckwith-Wiedemann syndrome. *Nat. Genet.* **14**, 171–173.

He, Y. and Ecker, J. R. (2015). Non-CG Methylation in the Human Genome. *Annu. Rev. Genomics Hum. Genet.* **16**, 55–77.

Heard, E. and Disteche, C. M. (2006). Dosage compensation in mammals: Fine-tuning the expression of the X chromosome. *Genes Dev.* **20**, 1848–1867.

Heard, E., Rougeulle, C., Arnaud, D., Avner, P., Allis, C. D. and Spector, D. L. (2001). Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell* **107**, 727–38.

Henckel, A., Chebli, K., Kota, S. K., Arnaud, P. and Feil, R. (2012). Transcription and histone methylation changes correlate with imprint acquisition in male germ cells. *EMBO J.* **31**, 606–615.

Henry, L.-A., Cassidy, T., McLaughlin, M., Pentieva, K., McNulty, H., Walsh, C. P. and Lees-Murdock, D. (2018). Folic Acid Supplementation throughout pregnancy: psychological developmental benefits for children. *Acta Paediatr.* **107**, 1370–1378.

Hernandez Mora, J. R., Tayama, C., Sánchez-Delgado, M., Monteagudo-Sánchez, A., Hata, K., Ogata, T., Medrano, J., Poo-Llanillo, M. E., Simón, C., Moran, S., et al. (2018). Characterization of parent-of-origin methylation using the Illumina Infinium MethylationEPIC array platform. *Epigenomics* **10**, 941–954.

Herzing, L. B. K., Kim, S.-J., Cook, E. H., Jr. and Ledbetter, D. H. (2001). The Human Aminophospholipid-Transporting ATPase Gene ATP10C Maps Adjacent to UBE3A and Exhibits Similar Imprinted Expression. *Am. J. Hum. Genet.* **68**, 1501.

Heyn, H., Li, N., Ferreira, H. J., Moran, S., Pisano, D. G., Gomez, A., Diez, J., Sanchez-Mut, J. V., Setien, F., Carmona, F. J., et al. (2012). Distinct DNA methylomes of newborns and centenarians. *Proc. Natl. Acad. Sci.* **109**, 10522–10527.

Heyn, P., Logan, C. V., Fluteau, A., Challis, R. C., Auchynnikava, T., Martin, C. A., Marsh, J. A., Taglini, F., Kilanowski, F., Parry, D. A., et al. (2019). Gain-of-function DNMT3A mutations cause microcephalic dwarfism and hypermethylation of Polycomb-regulated regions. *Nat. Genet.* **51**, 96–105.

Hikichi, T., Kohda, T., Kaneko-Ishino, T. and Ishino, F. (2003). Imprinting regulation of the murine Meg1/Grb10 and human GRB10 genes; roles of brain-specific promoters and mouse-specific CTCF-binding sites. *Nucleic Acids Res.* **31**, 1398–406.

Hiura, H., Okae, H., Miyauchi, N., Sato, F., Sato, A., Van De Pette, M., John, R. M., Kagami, M., Nakai, K., Soejima, H., et al. (2012). Characterization of DNA methylation errors in patients with imprinting disorders conceived by assisted reproduction technologies. *Hum. Reprod.* **27**, 2541–2548.

Hiura, H., Toyoda, M., Okae, H., Sakurai, M., Miyauchi, N., Sato, A., Kiyokawa, N., Okita, H., Miyagawa, Y., Akutsu, H., et al. (2013). Stability of genomic imprinting in human induced pluripotent stem cells. *BMC Genet.* **14**, 32.

Hjortshøj, T. D., Sørensen, A. R., Yusibova, M., Hansen, B. M., Dunø, M., Balslev-Harder, M., Grønskov, K., Hagen, J. M., Polstra, A. M., Eggermann, T., et al. (2020). upd(20)mat is a rare cause of the Silver-Russell-syndrome-like phenotype: Two unrelated cases and screening of large cohorts. *Clin. Genet.* **97**, 902–907.

Holliday, R. and Pugh, J. (1975). DNA modification mechanisms and gene activity during development. *Science (80-.).* **187**, 226–232.

Holt, L. J. and Siddle, K. (2005). Grb10 and Grb14: enigmatic regulators of insulin action – and more? *Biochem. J.* **388**, 393–406.

Horsthemke, B. and Buiting, K. (2006). Imprinting defects on human chromosome 15. *Cytogenet. Genome Res.* **113**, 292–299.

Horsthemke, B. and Wagstaff, J. (2008). Mechanisms of imprinting of the Prader-Willi/Angelman region. *Am. J. Med. Genet. Part A* **146**, 2041–2052.

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115.

Hoyo, C., Daltveit, A. K., Iversen, E., Benjamin-Neelon, S. E., Fuemmeler, B., Schildkraut, J., Murtha, A. P., Overcash, F., Vidal, A. C., Wang, F., et al. (2014). Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. *Epigenetics* **9**, 1120–1130.

Hulten, M. (1978). Selective Somatic Pairing and Fragility at 1q12 in a Boy with Common Variable Immuno Deficiency. *Clin. Genet.* **14**, 294–294.

Ibrahim, A., Kirby, G., Hardy, C., Dias, R. P., Tee, L., Lim, D., Berg, J., MacDonald, F., Nightingale, P. and Maher, E. R. (2014). Methylation analysis and diagnostics of Beckwith-Wiedemann syndrome in 1,000 subjects. *Clin. Epigenetics* **6**, 11.

Illingworth, R. S., Gruenewald-Schneider, U., Webb, S., Kerr, A. R. W., James, K. D., Turner, D. J., Smith, C., Harrison, D. J., Andrews, R. and Bird, A. P. (2010). Orphan CpG Islands Identify Numerous Conserved Promoters in the Mammalian Genome. *PLoS Genet.* **6**, e1001134.

Imbeault, M., Helleboid, P. Y. and Trono, D. (2017). KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature* **543**, 550–554.

Inoue, T., Yagasaki, H., Nishioka, J., Nakamura, A., Matsubara, K., Narumi, S., Nakabayashi, K., Yamazawa, K., Fuke, T., Oka, A., et al. (2018). Molecular and clinical analyses of two patients with UPD(16)mat detected by screening 94 patients with Silver-Russell

syndrome phenotype of unknown aetiology. *J. Med. Genet.* **56**, 413–418.

Ioannides, Y., Lokulo-Sodipe, K., Mackay, D. J. G., Davies, J. H. and Temple, I. K. (2014). Temple syndrome: improving the recognition of an underdiagnosed chromosome 14 imprinting disorder: an analysis of 51 published cases. *J. Med. Genet.* **51**, 495–501.

Irizarry, R. A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.* **41**, 178–186.

Irwin, R., Scullion, C., Thursby, S.-J., Sun, M., Thakur, A., Rothbart, S., Xu, G.-L. and Walsh, C. UHRF1 is required to suppress viral mimicry through both DNA methylation-dependent and -independent mechanisms.

Irwin, R. E., Thakur, A., O' Neill, K. M. and Walsh, C. P. (2014). 5-Hydroxymethylation marks a class of neuronal gene regulated by intragenic methylcytosine levels. *Genomics* **104**, 383–392.

Ishida, M. and Moore, G. E. (2013). The role of imprinted genes in humans. *Mol. Aspects Med.* **34**, 826–840.

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., He, C. and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–3.

Jang, H. S., Shin, W. J., Lee, J. E. and Do, J. T. (2017). CpG and non-CpG methylation in epigenetic gene regulation and brain function. *Genes (Basel).* **8**, 2–20.

Jasek, K., Kubatka, P., Samec, M., Liskova, A., Smejkal, K., Vybohova, D., Bugos, O., Biskupska-Bodova, K., Bielik, T., Zubor, P., et al. (2019). DNA methylation status in cancer disease: Modulations by plant-derived natural compounds and dietary

interventions. *Biomolecules* **9**,.

Jeffries, A. R., Uwanogho, D., Cocks, G., Perfect, L. W., Dempster, E., Mill, J. and Price, J. (2016). Erasure and reestablishment of random allelic expression imbalance after epigenetic reprogramming. *Rna* **22**, 1620–1630.

Jeffries, A. R., Maroofian, R., Salter, C. G., Chioza, B. A., Cross, H. E., Patton, M. A., Dempster, E., Karen Temple, I., Mackay, D. J. G., Rezwan, F. I., et al. (2019). Growth disrupting mutations in epigenetic regulatory molecules are associated with abnormalities of epigenetic aging. *Genome Res.* **29**, 1057–1066.

Jeltsch, A. and Jurkowska, R. Z. (2014). New concepts in DNA methylation. *Trends Biochem. Sci.* **39**, 310–318.

Jern, P. and Coffin, J. M. (2008). Effects of Retroviruses on Host Genome Function. *Annu. Rev. Genet.* **42**, 709–732.

Jiang, Y. L., Rigolet, M., Bourc'his, D., Nigon, F., Bokesoy, I., Fryns, J. P., Hultén, M., Jonveaux, P., Maraschio, P., Mégarbané, A., et al. (2005). DNMT3B mutations and DNA methylation defect define two types of ICF syndrome. *Hum. Mutat.* **25**, 56–63.

Jin, S. G., Jiang, Y., Qiu, R., Rauch, T. A., Wang, Y., Schackert, G., Krex, D., Lu, Q. and Pfeifer, G. P. (2011). 5-hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer Res.* **71**, 7360–7365.

Jjingo, D., Conley, A. B., Yi, S. V., Lunyak, V. V. and King Jordan, I. (2012). On the presence and role of human gene-body DNA methylation. *Oncotarget* **3**, 462–474.

John, R. M. and Lefebvre, L. (2011). Developmental regulation of somatic imprints. *Differentiation* **81**, 270–280.

Johnson, T. B. and Coghill, R. D. (1925). RESEARCHES ON PYRIMIDINES. C111. THE DISCOVERY

OF 5-METHYL-CYTOSINE IN TUBERCULINIC ACID, THE NUCLEIC ACID OF THE TUBERCLE BACILLUS [1]. *J. Am. Chem. Soc.* **47**, 2838–2844.

Johnson, M. H. and McConnell, J. M. L. (2004). Lineage allocation and cell polarity during mouse embryogenesis. *Semin. Cell Dev. Biol.* **15**, 583–597.

Jones, P. A. (1999). The DNA methylation paradox. *Trends Genet.* **15**, 34–7.

Jones, P. A. (2012). Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492.

Jordan, I. K., Rogozin, I. B., Glazko, G. V and Koonin, E. V (2003). Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet.* **19**, 68–72.

Joubert, B. R., Håberg, S. E., Nilsen, R. M., Wang, X., Vollset, S. E., Murphy, S. K., Huang, Z., Hoyo, C., Midttun, Ø., Cupul-Uicab, L. A., et al. (2012). 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ. Health Perspect.* **120**, 1425–31.

Jylhävä, J., Pedersen, N. L. and Hägg, S. (2017). Biological Age Predictors. *EBioMedicine* **21**, 29–36.

Kagami, M., O'Sullivan, M. J., Green, A. J., Watabe, Y., Arisaka, O., Masawa, N., Matsuoka, K., Fukami, M., Matsubara, K., Kato, F., et al. (2010). The IG-DMR and the MEG3-DMR at Human Chromosome 14q32.2: Hierarchical Interaction and Distinct Functional Properties as Imprinting Control Centers. *PLoS Genet.* **6**, e1000992.

Kagami, M., Nagasaki, K., Kosaki, R., Horikawa, R., Naiki, Y., Saitoh, S., Tajima, T., Yorifuji, T., Numakura, C., Mizuno, S., et al. (2017). Temple syndrome: Comprehensive molecular and clinical findings in 32 Japanese patients. *Genet. Med.* **19**, 1356–1366.

Kanduri, C. (2016). Long noncoding RNAs: Lessons from genomic imprinting. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1859**, 102–111.

Kappil, M., Lambertini, L. and Chen, J. (2015). Environmental Influences on Genomic Imprinting. *Curr. Environ. Heal. reports* **2**, 155–162.

Karagianni, P., Amazit, L., Qin, J. and Wong, J. (2008). ICBP90, a Novel Methyl K9 H3 Binding Protein Linking Protein Ubiquitination with Heterochromatin Formation. *Mol. Cell. Biol.* **28**, 705–717.

Karimi, M. M., Goyal, P., Maksakova, I. A., Bilenky, M., Leung, D., Tang, J. X., Shinkai, Y., Mager, D. L., Jones, S., Hirst, M., et al. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mescs. *Cell Stem Cell* **8**, 676–687.

Kaufman, M. H., Robertson, E. J., Handyside, A. H. and Evans, M. J. (1983). Establishment of pluripotential cell lines from haploid mouse embryos. *J. Embryol. Exp. Morphol.* **Vol. 73**, 249–261.

Kernohan, K. D., Cigana Schenkel, L., Huang, L., Smith, A., Pare, G., Ainsworth, P., Care4Rare Canada Consortium, Boycott, K. M., Warman-Chardon, J. and Sadikovic, B. (2016). Identification of a methylation profile for DNMT1-associated autosomal dominant cerebellar ataxia, deafness, and narcolepsy. *Clin. Epigenetics* **8**,.

Kim, K., Ng, K., Rugg-Gunn, P. J., Shieh, J. H., Kirak, O., Jaenisch, R., Wakayama, T., Moore, M. A., Pedersen, R. A. and Daley, G. Q. (2007). Recombination Signatures Distinguish Embryonic Stem Cells Derived by Parthenogenesis and Somatic Cell Nuclear Transfer. *Cell Stem Cell* **1**, 346–352.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M. J., Ji, H., Ehrlich, L. I. R., et al. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285–290.

Klein, C. J., Botuyan, M. V., Wu, Y., Ward, C. J., Nicholson, G. A., Hammans, S., Hojo, K., Yamanishi, H., Karpf, A. R., Wallace, D. C., et al. (2011). Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss. *Nat. Genet.* **43**, 595–600.

Klose, R. J. and Bird, A. P. (2006). Genomic DNA methylation: The mark and its mediators. *Trends Biochem. Sci.* **31**, 89–97.

Ko, M. S. H. (2001). Embryogenomics: Developmental biology meets genomics. *Trends Biotechnol.* **19**, 511–518.

Kojima, Y., Tam, O. H. and Tam, P. P. L. (2014). Timing of developmental events in the early mouse embryo. *Semin. Cell Dev. Biol.* **34**, 65–75.

Kong, X., Chen, J., Xie, W., Brown, S. M., Cai, Y., Wu, K., Fan, D., Nie, Y., Yegnasubramanian, S., Tiedemann, R. L., et al. (2019). Defining UHRF1 Domains that Support Maintenance of Human Colon Cancer DNA Methylation and Oncogenic Properties. *Cancer Cell* **35**, 633-648.e7.

Kozlenkov, A., Roussos, P., Timashpolsky, A., Barbu, M., Rudchenko, S., Bibikova, M., Klotzle, B., Byne, W., Lyddon, R., Di Narzo, A. F., et al. (2014). Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites. *Nucleic Acids Res.* **42**, 109–127.

Kuchmiy, A. A., D'Hont, J., Hochepied, T. and Lamkanfi, M. (2016). NLRP2 controls age-associated maternal fertility. *J. Exp. Med.* **213**, 2851–2860.

Kurihara, Y., Kawamura, Y., Uchijima, Y., Amamo, T., Kobayashi, H., Asano, T. and Kurihara, H. (2008). Maintenance of genomic methylation patterns during preimplantation development requires the somatic form of DNA methyltransferase 1. *Dev. Biol.* **313**, 335–346.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K.,

Dewar, K., Doyle, M., Fitzhugh, W., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.

Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J. and Reik, W. (2003). Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35**, 88–93.

Larsen, F., Gundersen, G., Lopez, R. and Prydz, H. (1992). CpG islands as gene markers in the human genome. *Genomics* **13**, 1095–107.

Lees-Murdock, D. J., De Felici, M. and Walsh, C. P. (2003). Methylation dynamics of repetitive DNA elements in the mouse germ cell lineage. *Genomics* **82**, 230–237.

Leff, S. E., Brannan, C. I., Reed, M. L., Oz, T., Francke, U., Copeland, N. G. and Jenkins, N. A. (1992). *Maternal imprinting of the mouse Snrpn gene and conserved linkage homology with the human Prader-Willi syndrome region*.

Lei, H., Oh, S. P. P., Okano, M., Jüttermann, R., Goss, K. A. A., Jaenisch, R., Li, E., Juttermann, R., Goss, K. A. A., Jaenisch, R., et al. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. **122**, 3195–205.

Leonhardt, H., Page, A. W., Weier, H. U. and Bestor, T. H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71**, 865–73.

Leseva, M., Santostefano, K. E., Rosenbluth, A. L., Hamazaki, T. and Terada, N. (2013). E2f6-mediated repression of the meiotic Stag3 and Smc1β genes during early embryonic development requires Ezh2 and not the de novo methyltransferase Dnmt3b. *Epigenetics* **8**, 873–884.

Lewis, M. W., Brant, J. O., Kramer, J. M., Moss, J. I., Yang, T. P., Hansen, P. J., Williams, R. S. and Resnick, J. L. (2015). Angelman syndrome imprinting center encodes a transcriptional promoter. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6871–6875.

Ley, T. J., Ding, L., Walter, M. J., McLellan, M. D., Lamprecht, T., Larson, D. E., Kandoth, C., Payton, J. E., Baty, J., Welch, J., et al. (2010). DNMT3A mutations in acute myeloid leukemia. *N. Engl. J. Med.* **363**, 2424–2433.

Li, E., Bestor, T. H. and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926.

Li, E., Beard, C. and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature* **366**, 362–365.

Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P. and Ferguson-Smith, A. C. (2008). A Maternal-Zygotic Effect Gene, Zfp57, Maintains Both Maternal and Paternal Imprints. *Dev. Cell* **15**, 547–557.

Li, T., Wang, L., Du, Y., Xie, S., Yang, X., Lian, F., Zhou, Z. and Qian, C. (2018). Structural and mechanistic insights into UHRF1-mediated DNMT1 activation in the maintenance DNA methylation. *Nucleic Acids Res.* **46**, 3218–3231.

Liao, J., Karnik, R., Gu, H., Ziller, M. J., Clement, K., Tsankov, A. M., Akopian, V., Gifford, C. A., Donaghey, J., Galonska, C., et al. (2015). Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat. Genet.* **47**, 469–78.

Lim, D. H. K. and Maher, E. R. (2010). *Genomic imprinting syndromes and cancer*.

Lin, M. E., Hou, H. A., Tsai, C. H., Wu, S. J., Kuo, Y. Y., Tseng, M. H., Liu, M. C., Liu, C. W., Chou, W. C., Chen, C. Y., et al. (2018). Dynamics of DNMT3A mutation and prognostic relevance in patients with primary myelodysplastic syndrome. *Clin. Epigenetics* **10**,.

Lindholm, M. E., Marabita, F., Gomez-Cabrero, D., Rundqvist, H., Ekström, T. J., Tegnér, J. and Sundberg, C. J. (2014). An integrative analysis reveals coordinated reprogramming of the epigenome and the transcriptome in human skeletal muscle after training. *Epigenetics* **9**, 1557.

Ling, C., Del Guerra, S., Lupi, R., Rönn, T., Granhall, C., Luthman, H., Masiello, P., Marchetti, P., Groop, L. and Del Prato, S. (2008). Epigenetic regulation of PPARGC1A in human type 2 diabetic islets and effect on insulin secretion. *Diabetologia* **51**, 615–22.

Lio, C. W. J. and Rao, A. (2019). TET enzymes and 5hMC in adaptive and innate immune systems. *Front. Immunol.* **10**,.

Lock, L. F., Melton, D. W., Caskey, C. T. and Martin, G. R. (1986). Methylation of the mouse hprt gene differs on the active and inactive X chromosomes. *Mol. Cell. Biol.* **6**, 914–924.

Loughery, J. E. ., Dunne, P. D., O'Neill, K. M., Meehan, R. R., McDaid, J. R. and Walsh, C. P. (2011). DNMT1 deficiency triggers mismatch repair defects in human cells through depletion of repair protein levels in a process involving the DNA damage response. *Hum. Mol. Genet.* **20**, 3241–3255.

Luedi, P. P., Dietrich, F. S., Weidman, J. R., Bosko, J. M., Jirtle, R. L. and Hartemink, A. J. (2007). Computational and experimental identification of novel human imprinted genes. *Genome Res.* **17**, 1723–1730.

Luk, H. M. (2016). Temple syndrome misdiagnosed as Silver-Russell syndrome. *Clin. Dysmorphol.* **25**, 82–83.

Luk, H. M. (2017). Clinical and molecular characterization of Beckwith-Wiedemann syndrome in a Chinese population. *J. Pediatr. Endocrinol. Metab.* **30**, 89–95.

LYON, M. F. (1961). Gene Action in the X-chromosome of the Mouse (Mus musculus L.). *Nature* **190**, 372–373.

Maatouk, D. M., Kellam, L. D., Mann, M. R. W., Lei, H., Li, E., Bartolomei, M. S. and Resnick, J. L. (2006). DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* **133**, 3411–3418.

Mackay, D. J. G., Temple, I. K., Shield, J. P. H. and Robinson, D. O. (2005). Bisulphite sequencing of the transient neonatal diabetes mellitus DMR facilitates a novel diagnostic test but reveals no methylation anomalies in patients of unknown aetiology. *Hum. Genet.* **116**, 255–261.

Mackay, D. J. G., Callaway, J. L. A., Marks, S. M., White, H. E., Acerini, C. L., Boonen, S. E., Dayanikli, P., Firth, H. V., Goodship, J. A., Haemers, A. P., et al. (2008). Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. *Nat. Genet.* **40**, 949–951.

Maeda, T., Higashimoto, K., Jozaki, K., Yatsuki, H., Nakabayashi, K., Makita, Y., Tonoki, H., Okamoto, N., Takada, F., Ohashi, H., et al. (2014). Comprehensive and quantitative multilocus methylation analysis reveals the susceptibility of specific imprinted differentially methylated regions to aberrant methylation in Beckwith-Wiedemann syndrome with epimutations. *Genet. Med.* **16**, 903–912.

Maenohara, S., Unoki, M., Toh, H., Ohishi, H., Sharif, J., Koseki, H. and Sasaki, H. (2017). Role of UHRF1 in de novo DNA methylation in oocytes and maintenance methylation in preimplantation embryos. *PLoS Genet.* **13**,.

Mahadevan, S., Wen, S., Wan, Y. W., Peng, H. H., Otta, S., Liu, Z., Iacovino, M., Mahen, E. M., Kyba, M., Sadikovic, B., et al. (2014). NLRP7 affects trophoblast lineage differentiation, binds to overexpressed YY1 and alters cpg methylation. *Hum. Mol. Genet.* **23**, 706–716.

Mahadevan, S., Sathappan, V., Utama, B., Lorenzo, I., Kaskar, K. and Van den Veyver, I. B. (2017). Maternally expressed NLRP2 links the subcortical maternal complex (SCMC) to fertility, embryogenesis and epigenetic reprogramming. *Sci. Rep.* **7**, 44667.

Maiti, A. and Drohat, A. C. (2011). Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine. *J. Biol. Chem.* **286**, 35334–35338.

Mandal, P. K. and Kazazian, H. H. (2008). SnapShot: Vertebrate Transposons. *Cell* **135**,.

Mann, J. R., Gadi, I., Harbison, M. L., Abbondanzo, S. J. and Stewart, C. L. (1990). Androgenetic mouse embryonic stem cells are pluripotent and cause skeletal defects in chimeras: Implications for genetic imprinting. *Cell* **62**, 251–260.

Maranga, C., Fernandes, T. G., Bekman, E. and da Rocha, S. T. (2020). Angelman syndrome: a journey through the brain. *FEBS J.* **287**, 2154–2175.

Marioni, R. E., Shah, S., McRae, A. F., Chen, B. H., Colicino, E., Harris, S. E., Gibson, J., Henders, A. K., Redmond, P., Cox, S. R., et al. (2015). DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol.* **16**, 25.

Markunas, C. A., Xu, Z., Harlid, S., Wade, P. A., Lie, R. T., Taylor, J. A. and Wilcox, A. J. (2014). Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. *Environ. Health Perspect.* **122**, 1147–1153.

Martin, S. L., Cruceanu, M., Branciforte, D., Li, P. W. L., Kwok, S. C., Hodges, R. S. and Williams, M. C. (2005). LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 Protein. *J. Mol. Biol.* **348**, 549–561.

Mathias, S. L., Scott, A. F., Kazazian, H. H., Boeke, J. D. and Gabriel, A. (1991). Reverse transcriptase encoded by a human transposable element. *Science (80-.).* **254**, 1808–1810.

Matsubara, K., Itoh, M., Shimizu, K., Saito, S., Enomoto, K., Nakabayashi, K., Hata, K., Kurosawa, K., Ogata, T., Fukami, M., et al. (2019). Exploring the unique function of imprinting control centers in the PWS/AS-responsible region: Finding from array-based methylation analysis in cases with variously sized microdeletions. *Clin. Epigenetics* **11**,.

Matsui, T., Leung, D., Miyashita, H., Maksakova, I. A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M. C. and Shinkai, Y. (2010). Proviral silencing in embryonic stem cells

requires the histone methyltransferase ESET. *Nature* **464**, 927–931.

Maunakea, A. K., Nagarajan, R. P., Bilenky, M., Ballinger, T. J., Dsouza, C., Fouse, S. D., Johnson, B. E., Hong, C., Nielsen, C., Zhao, Y., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466**, 253–257.

McCann, J. A., Zheng, H., Islam, A., Goodyer, C. G. and Polychronakos, C. (2001). Evidence against GRB10 as the gene responsible for Silver-Russell syndrome. *Biochem. Biophys. Res. Commun.* **286**, 943–948.

McClelland, R., Christensen, K., Mohammed, S., McGuinness, D., Cooney, J., Bakshi, A., Demou, E., MacDonald, E., Caslake, M., Stenvinkel, P., et al. (2016). Accelerated ageing and renal dysfunction links lower socioeconomic status and dietary phosphate intake. *Aging (Albany. NY).* **8**, 1135–49.

McEwen, K. R. and Ferguson-Smith, A. C. (2010). Distinguishing epigenetic marks of developmental and imprinting regulation. *Epigenetics and Chromatin* **3**,.

McGrath, J. and Solter, D. (1983). Nuclear transplantation in mouse embryos. *J. Exp. Zool.* **228**, 355–362.

Meilinger, D., Fellinger, K., Bultmann, S., Rothbauer, U., Bonapace, I. M., Klinkert, W. E. F., Spada, F. and Leonhardt, H. (2009). Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. *EMBO Rep.* **10**, 1259–1264.

Meissner, A., Mikkelsen, T. S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B. E., Nusbaum, C., Jaffe, D. B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770.

Meyer, E., Lim, D., Pasha, S., Tee, L. J., Rahman, F., Yates, J. R. W., Woods, C. G., Reik, W. and Maher, E. R. (2009). Germline Mutation in NLRP2 (NALP2) in a Familial Imprinting

Disorder (Beckwith-Wiedemann Syndrome). *PLoS Genet.* **5**, e1000423.

Mihajlović, A. I. and Bruce, A. W. (2017). The first cell-fate decision of mouse preimplantation embryo development: Integrating cell position and polarity. *Open Biol.* **7**,.

Mikkelsen, T. S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B. E., Jaenisch, R., Lander, E. S. and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49–55.

Min, H.-Y., Lee, S.-C., Woo, J. K., Jung, H. J., Park, K. H., Jeong, H. M., Hyun, S. Y., Cho, J., Lee, W., Park, J. E., et al. (2017). Essential Role of DNA Methyltransferase 1-mediated Transcription of Insulin-like Growth Factor 2 in Resistance to Histone Deacetylase Inhibitors. *Clin. Cancer Res.* **23**, 1299–1311.

Mitiku, N. and Baker, J. C. (2007). Genomic Analysis of Gastrulation and Organogenesis in the Mouse. *Dev. Cell* **13**, 897–907.

Moarefi, A. H. and Chédin, F. (2011). ICF syndrome mutations cause a broad spectrum of biochemical defects in DNMT3B-mediated de novo DNA methylation. *J. Mol. Biol.* **409**, 758–772.

Monk, D., Arnaud, P., Frost, J., Hills, F. A., Stanier, P., Feil, R. and Moore, G. E. (2009). Reciprocal imprinting of human GRB10 in placental trophoblast and brain: evolutionary conservation of reversed allelic expression. *Hum. Mol. Genet.* **18**, 3066–3074.

Monk, D., Sanchez-Delgado, M. and Fisher, R. (2017). NLRPS, the subcortical maternal complex and genomic imprinting. *Reproduction* **154**, R161–R170.

Monk, D., Morales, J., den Dunnen, J. T., Russo, S., Court, F., Prawitt, D., Eggermann, T., Beygo, J., Buiting, K. and Tümer, Z. (2018). Recommendations for a nomenclature system for reporting methylation aberrations in imprinted domains. *Epigenetics* **13**, 117–121.

Monk, D., Mackay, D. J. G., Eggermann, T., Maher, E. R. and Riccio, A. (2019). Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. *Nat. Rev. Genet.* **20**, 235–248.

Monteagudo-Sánchez, A., Sánchez-Delgado, M., Mora, J. R. H., Santamaría, N. T., Gratacós, E., Esteller, M., De Heredia, M. L., Nunes, V., Choux, C., Fauque, P., et al. (2019). Differences in expression rather than methylation at placenta-specific imprinted loci is associated with intrauterine growth restriction. *Clin. Epigenetics* **11**, 35.

Monteagudo-Sánchez, A., Hernandez Mora, J. R., Simon, C., Burton, A., Tenorio, J., Lapunzina, P., Clark, S., Esteller, M., Kelsey, G., López-Siguero, J. P., et al. (2020). The role of ZFP57 and additional KRAB-zinc finger proteins in the maintenance of human imprinted methylation and multi-locus imprinting disturbances. *Nucleic Acids Res.* **48**, 11394–11407.

Morales, C., Soler, A., Badenas, C., Rodríguez-Revenga, L., Nadal, A., Martínez, J. M., Mademont-Soler, I., Borrell, A., Milà, M. and Sánchez, A. (2009). Reproductive consequences of genome-wide paternal uniparental disomy mosaicism: description of two cases with different mechanisms of origin and pregnancy outcomes. *Fertil. Steril.* **92**, 393.e5-393.e9.

Mousli, M., Hopfner, R., Abbady, A. Q., Monté, D., Jeanblanc, M., Oudet, P., Louis, B. and Bronner, C. (2003). ICBP90 belongs to a new family of proteins with an expression that is deregulated in cancer cells. *Br. J. Cancer* **89**, 120–127.

Murdoch, S., Djuric, U., Mazhar, B., Seoud, M., Khan, R., Kuick, R., Bagga, R., Kircheisen, R., Ao, A., Ratti, B., et al. (2006). Mutations in NALP7 cause recurrent hydatidiform moles and reproductive wastage in humans. *Nat. Genet.* **38**, 300–302.

Murrell, A., Heeson, S., Bowden, L., Constância, M., Dean, W., Kelsey, G. and Reik, W. (2001). An intragenic methylated region in the imprinted Igf2 gene augments transcription. *EMBO Rep.* **2**, 1101–1106.

Muto, M., Kanari, Y., Kubo, E., Takabe, T., Kurihara, T., Fujimori, A. and Tatsumi, K. (2002). Targeted disruption of Np95 gene renders murine embryonic stem cells hypersensitive to DNA damaging agents and DNA replication blocks. *J. Biol. Chem.* **277**, 34549–34555.

Nakaya, Y., Sukowati, E. W. and Sheng, G. (2013). Epiblast integrity requires CLASP and Dystroglycan-mediated microtubule anchoring to the basal cortex. *J. Cell Biol.* **202**, 637–651.

Nechin, J., Tunstall, E., Raymond, N., Hamagami, N., Pathmanabhan, C., Forestier, S. and Davis, T. L. (2019). Hemimethylation of CpG dyads is characteristic of secondary DMRs associated with imprinted loci and correlates with 5-hydroxymethylcytosine at paternally methylated sequences. *Epigenetics Chromatin* **12**, 64.

Nestor, C. E., Ottaviano, R., Reddington, J., Sproul, D., Reinhardt, D., Dunican, D., Katz, E., Dixon, J. M., Harrison, D. J. and Meehan, R. R. (2012). Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res.* **22**, 467–477.

Nicholls, R. D. and Knepper, J. L. (2001). Genome Organization, Function, and Imprinting in Prader-Willi and Angelman Syndromes. *Annu. Rev. Genomics Hum. Genet.* **2**, 153–175.

Nikaido, I., Saito, C., Mizuno, Y., Meguro, M., Bono, H., Kadomura, M., Kono, T., Morris, G. A., Lyons, P. A., Oshimura, M., et al. (2003). Discovery of imprinted transcripts in the mouse transcriptome using large-scale expression profiling. *Genome Res.* **13**, 1402–1409.

Nishino, K., Toyoda, M., Yamazaki-Inoue, M., Fukawatase, Y., Chikazawa, E., Sakaguchi, H., Akutsu, H. and Umezawa, A. (2011). DNA methylation dynamics in human induced pluripotent stem cells over time. *PLoS Genet.* **7**, e1002085.

Norvil, A. B., Saha, D., Dar, M. S. and Gowher, H. (2019). Effect of disease-associated germline mutations on structure function relationship of DNA methyltransferases. *Genes (Basel).* **10**,.

Nygren, A. O. H., Ameziane, N., Duarte, H. M. B., Vijzelaar, R. N. C. P., Waisfisz, Q., Hess, C. J., Schouten, J. P. and Errami, A. (2005). Methylation-Specific MLPA (MS-MLPA): Simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. *Nucleic Acids Res.* **33**, 1–9.

O'Brien, E. K. and Wolf, J. B. (2017). The coadaptation theory for genomic imprinting. *Evol. Lett.* **1**, 49–59.

O'Neill, K. M., Irwin, R. E., Mackin, S.-J., Thursby, S.-J., Thakur, A., Bertens, C., Masala, L., Loughery, J. E. P., McArt, D. G. and Walsh, C. P. (2018). Depletion of DNMT1 in differentiated human cells highlights key classes of sensitive genes and an interplay with polycomb repression. *Epigenetics Chromatin* **11**, 12.

Ogata, T. and Kagami, M. (2016). Kagami-Ogata syndrome: A clinically recognizable upd(14)pat and related disorder affecting the chromosome 14q32.2 imprinted region. *J. Hum. Genet.* **61**, 87–94.

Ohi, Y., Qin, H., Hong, C., Blouin, L., Polo, J. M., Guo, T., Qi, Z., Downey, S. L., Manos, P. D., Rossi, D. J., et al. (2011). Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nat. Cell Biol.* **13**, 541–549.

Ohtani, H., Liu, M., Zhou, W., Liang, G. and Jones, P. A. (2018). Switching roles for DNA and histone methylation depend on evolutionary ages of human endogenous retroviruses. *Genome Res.* **28**, 1147–1157.

Okano, M., Xie, S. and Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* **19**, 219–220.

Okano, M., Bell, D. W., Haber, D. A. and Li, E. (1999). DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell* **99**, 247–257.

Ondičová, M., Oakey, R. J. and Walsh, C. P. (2020). Is imprinting the result of “friendly fire” by the host defense system? *PLoS Genet.* **16**,.

Ooi, S. K. T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.-P., Allis, C. D., et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714–717.

Ordoñez, R., Martínez-Calle, N., Agirre, X. and Prosper, F. (2019). DNA methylation of enhancer elements in myeloid neoplasms: Think outside the promoters? *Cancers (Basel).* **11**,.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W. and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10**, 475–478.

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M. and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep.* **10**, 1235–1241.

Ouellette, M. M. (2000). The establishment of telomerase-immortalized cell lines representing human chromosome instability syndromes. *Hum. Mol. Genet.* **9**, 403–411.

Õunap, K. (2016). Silver-Russell Syndrome and Beckwith-Wiedemann Syndrome: Opposite Phenotypes with Heterogeneous Molecular Etiology. *Mol. Syndromol.* **7**, 110–121.

Pacaud, R., Brocard, E., Lalier, L., Hervouet, E., Vallette, F. M. and Cartron, P. F. (2014). The DNMT1/PCNA/UHRF1 disruption induces tumorigenesis characterized by similar genetic and epigenetic signatures. *Sci. Rep.* **4**, 1–9.

Panning, B. and Jaenisch, R. (1996). DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev.* **10**, 1991–2002.

Papale, L. A., Madrid, A., Li, S. and Alisch, R. S. (2017). Early-life stress links 5-hydroxymethylcytosine to anxiety-related behaviors. *Epigenetics* **12**, 264–276.

Paroush, Z., Keshet, I., Yisraeli, J. and Cedar, H. (1990). Dynamics of demethylation and activation of the α-actin gene in myoblasts. *Cell* **63**, 1229–1237.

Paulsen, M. and Ferguson-Smith, A. C. (2001). DNA methylation in genomic imprinting, development, and disease. *J. Pathol.* **195**, 97–110.

Pawlak, M. and Jaenisch, R. (2011). De novo DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state. *Genes Dev.* **25**, 1035–1040.

Pervjakova, N., Kasela, S., Morris, A. P., Kals, M., Metspalu, A., Lindgren, C. M., Salumets, A. and Mägi, R. (2016). Imprinted genes and imprinting control regions show predominant intermediate methylation in adult somatic tissues. *Epigenomics* **8**, 789–799.

Peters, J. (2014). The role of genomic imprinting in biology and disease: An expanding view. *Nat. Rev. Genet.* **15**, 517–530.

Petit, F., Holder-Espinasse, M., Duban-Bedu, B., Bouquillon, S., Boute-Benejean, O., Bazin, A., Rouland, V., Manouvrier-Hanu, S. and Delobel, B. (2012). Trisomy 7 mosaicism prenatally misdiagnosed and maternal uniparental disomy in a child with pigmentary mosaicism and Russell- Silver syndrome. *Clin. Genet.* **81**, 265–71.

Petre, G., Lorès, P., Sartelet, H., Truffot, A., Poreau, B., Brandeis, S., Martinez, G., Satre, V., Harbuz, R., Ray, P. F., et al. (2018). Genomic duplication in the 19q13.42 imprinted region identified as a new genetic cause of intrauterine growth restriction. *Clin. Genet.* **94**, 575–580.

Pidsley, R., Y Wong, C. C., Volta, M., Lunnon, K., Mill, J. and Schalkwyk, L. C. (2013). A data-

driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics* **14**, 293.

Pilvar, D., Reiman, M., Pilvar, A. and Laan, M. (2019). Parent-of-origin-specific allelic expression in the human placenta is limited to established imprinted loci and it is stably maintained across pregnancy. *Clin. Epigenetics* **11**, 1–14.

Pinheiro, I. and Heard, E. (2017). X chromosome inactivation: new players in the initiation of gene silencing. *F1000Research* **6**,.

Píškala, A. and Šorm, F. (1964). Nucleic acids components and their analogues. LI. Synthesis of 1-glycosyl derivatives of 5-azauracil and 5-azacytosine. *Collect. Czechoslov. Chem. Commun.* **29**, 2060–2076.

Plasschaert, R. N. and Bartolomei, M. S. (2015). Tissue-specific regulation and function of Grb10 during growth and neuronal commitment. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6841–7.

Plasterk, R. H. A., Izsvák, Z. and Ivics, Z. (1999). Resident aliens the Tc1/mariner superfamily of transposable elements. *Trends Genet.* **15**, 326–332.

Polepalli, S., George, S. M., Valli Sri Vidya, R., Rodrigues, G. S., Ramachandra, L., Chandrashekar, R., M, D. N., Rao, P. P. N., Pestell, R. G. and Rao, M. (2019). Role of UHRF1 in malignancy and its function as a therapeutic target for molecular docking towards the SRA domain. *Int. J. Biochem. Cell Biol.* **114**, 105558.

Poole, R. L., Docherty, L. E., Al Sayegh, A., Caliebe, A., Turner, C., Baple, E., Wakeling, E., Harrison, L., Lehmann, A., Temple, I. K., et al. (2013). Targeted methylation testing of a patient cohort broadens the epigenetic and clinical description of imprinting disorders. *Am. J. Med. Genet. Part A* **161**, 2174–2182.

Posynick, B. J. and Brown, C. J. (2019). Escape From X-Chromosome Inactivation: An

Evolutionary Perspective. *Front. Cell Dev. Biol.* **7**,.

Prader, A. (1956). Ein Syndrom von Adipositas, Kleinwuchs, Kryptorchismus und Oligophrenie nach myatonieartigem Zustand im Neugeborenenalter. *Schweiz Med Wochenschr* **86**, 1260–1261.

Pradhan, M., Estève, P. O., Hang, G. C., Samaranayke, M., Kim, G. Do and Pradhan, S. (2008). CXXC domain of human DNMT1 is essential for enzymatic activity. *Biochemistry* **47**, 10000–10009.

Prasasya, R., Grotheer, K. V, Siracusa, linda D. and Bartolomei, M. S. (2020). Temple Syndrome and Kagami-Ogata Syndrome: Clinical Presentations, Genotypes, Models and Mechanisms . *Hum. Mol. Genet.*

Prickett, A. R., Ishida, M., Böhm, S., Frost, J. M., Puszyk, W., Abu-Amero, S., Stanier, P., Schulz, R., Moore, G. E. and Oakey, R. J. (2015). Genome-wide methylation analysis in Silver-Russell syndrome patients. *Hum. Genet.* **134**, 317–332.

Proudhon, C., Duffié, R., Ajjan, S., Cowley, M., Iranzo, J., Carbajosa, G., Saadeh, H., Holland, M. L., Oakey, R. J., Rakyan, V. K., et al. (2012). Protection against de novo methylation is instrumental in maintaining parent-of-origin methylation inherited from the gametes. *Mol. Cell* **47**, 909–20.

Qi, S., Wang, Z., Li, P., Wu, Q., Shi, T., Li, J. and Wong, J. (2015). Non-germ line restoration of genomic imprinting for a small subset of imprinted genes in ubiquitin-like PHD and RING finger domain-containing 1 (Uhrf1) null mouse embryonic stem cells. *J. Biol. Chem.* **290**, 14181–14191.

Qian, X., Ba, Y., Zhuang, Q. and Zhong, G. (2014). RNA-seq technology and its application in fish transcriptomics. *Omi. A J. Integr. Biol.* **18**, 98–110.

Qian, J., Nguyen, N. M. P., Rezaei, M., Huang, B., Tao, Y., Zhang, X., Cheng, Q., Yang, H.,

Asangla, A., Majewski, J., et al. (2018). Biallelic PADI6 variants linking infertility, miscarriages, and hydatidiform moles. *Eur. J. Hum. Genet.* **26**, 1007–1013.

Qu, Y., Lennartsson, A., Gaidzik, V. I., Deneberg, S., Karimi, M., Bengtzén, S., Höglund, M., Bullinger, L., Döhner, K. and Lehmann, S. (2014). Differential methylation in CN-AML preferentially targets non-CGI regions and is dictated by DNMT3A mutational status and associated with predominant hypomethylation of HOX genes. *Epigenetics* **9**, 1108–1119.

Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P. V., Grimaldi, G., Riccio, A., et al. (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol. Cell* **44**, 361–372.

Quivoron, C., Couronné, L., Della Valle, V., Lopez, C. K., Plo, I., Wagner-Ballon, O., Do Cruzeiro, M., Delhommeau, F., Arnulf, B., Stern, M. H., et al. (2011). TET2 Inactivation Results in Pleiotropic Hematopoietic Abnormalities in Mouse and Is a Recurrent Event during Human Lymphomagenesis. *Cancer Cell* **20**, 25–38.

Rabinovitz, S., Kaufman, Y., Ludwig, G., Razin, A. and Shemer, R. (2012). Mechanisms of activation of the paternally expressed genes by the Prader-Willi imprinting center in the Prader-Willi/Angelman syndromes domains. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 7403–7408.

Ramsahoye, B. H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A. P. and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 5237–5242.

Rasmussen, K. D. and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* **30**, 733–750.

Reik, W., Dean, W. and Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science (80-.).* **293**, 1089–1093.

Renfree, M. B., Suzuki, S. and Kaneko-Ishino, T. (2013). The origin and evolution of genomic imprinting and viviparity in mammals. *Philos. Trans. R. Soc. B Biol. Sci.* **368**,.

Revazova, E. S., Turovets, N. A., Kochetkova, O. D., Kindarova, L. B., Kuzmichev, L. N., Janus, J. D. and Pryzhkova, M. V. (2007). Patient-specific stem cell lines derived from human parthenogenetic blastocysts. *Cloning Stem Cells* **9**, 432–449.

Rezwan, F. I., Docherty, L. E., Poole, R. L., Lockett, G. A., Arshad, S. H., Holloway, J. W., Temple, I. K. and Mackay, D. J. (2015). A statistical method for single sample analysis of humanmethylation450 array data: Genome-wide methylation analysis of patients with imprinting disorders. *Clin. Epigenetics* **7**,.

Rhee, I., Jair, K.-W., Yen, R.-W. C., Lengauer, C., Herman, J. G., Kinzler, K. W., Vogelstein, B., Baylin, S. B. and Schuebel, K. E. (2000). CpG methylation is maintained in human cancer cells lacking DNMT1. *Nature* **404**, 1003–1007.

Rhee, I., Bachman, K. E., Park, B. H., Jair, K.-W., Yen, R.-W. C., Schuebel, K. E., Cui, H., Feinberg, A. P., Lengauer, C., Kinzler, K. W., et al. (2002). DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature* **416**, 552–556.

Riordan, J. D., Keng, V. W., Tschida, B. R., Scheetz, T. E., Bell, J. B., Podetz-Pedersen, K. M., Moser, C. D., Copeland, N. G., Jenkins, N. A., Roberts, L. R., et al. (2013). Identification of Rtl1, a Retrotransposon-Derived Imprinted Gene, as a Novel Driver of Hepatocarcinogenesis. *PLoS Genet.* **9**,.

Rochtus, A., Martin-Trujillo, A., Izzi, B., Elli, F., Garin, I., Linglart, A., Mantovani, G., Perez de Nanclares, G., Thiele, S., Decallonne, B., et al. (2016). Genome-wide DNA methylation analysis of pseudohypoparathyroidism patients with GNAS imprinting defects. *Clin. Epigenetics* **8**, 1–12.

Romanelli, V., Nevado, J., Fraga, M., Trujillo, A. M., Mori, M. Á., Fernández, L., De Nanclares, G. P., Martínez-Glez, V., Pita, G., Meneses, H., et al. (2011). Constitutional mosaic

genome-wide uniparental disomy due to diploidisation: An unusual cancer-predisposing mechanism. *J. Med. Genet.* **48**, 212–216.

Rönnerblad, M., Andersson, R., Olofsson, T., Douagi, I., Karimi, M., Lehmann, S., Hoof, I., De Hoon, M., Itoh, M., Nagao-Sato, S., et al. (2014). Analysis of the DNA methylome and transcriptome in granulopoiesis reveals timed changes and dynamic enhancer methylation. *Blood* **123**,.

Rothbart, S. B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A. I., Barsyte-Lovejoy, D., Martinez, J. Y., Bedford, M. T., Fuchs, S. M., et al. (2012). Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat. Struct. Mol. Biol.* **19**, 1155–1160.

Rougeulle, C., Glatt, H. and Lalande, M. (1997). The Angelman syndrome candidate gene, UBE3AIE6-AP, is imprinted in brain. *Nat. Genet.* **17**, 14–15.

Rougier, N., Bourc'his, D., Gomes, D. M., Niveleau, A., Plachot, M., Pàldi, A. and Viegas-Péquignot, E. (1998). Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev.* **12**, 2108–13.

Roulois, D., Loo Yau, H., Singhania, R., Wang, Y., Danesh, A., Shen, S. Y., Han, H., Liang, G., Jones, P. A., Pugh, T. J., et al. (2015). DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. *Cell* **162**, 961–973.

Russell, A. (1954). A syndrome of intra-uterine dwarfism recognizable at birth with cranio-facial dysostosis, disproportionately short arms, and other anomalies (5 examples). *Proc. R. Soc. Med.* **47**, 1040–4.

Russler-Germain, D. A., Spencer, D. H., Young, M. A., Lamprecht, T. L., Miller, C. A., Fulton, R., Meyer, M. R., Erdmann-Gilmore, P., Townsend, R. R., Wilson, R. K., et al. (2014). The R882H DNMT3A Mutation Associated with AML Dominantly Inhibits Wild-Type DNMT3A by Blocking Its Ability to Form Active Tetramers. *Cancer Cell* **25**, 442–454.

Rutledge, C. E., Thakur, A., O'Neill, K. M., Irwin, R. E., Sato, S., Hata, K. and Walsh, C. P. (2014). Ontogeny, conservation and functional significance of maternally inherited DNA methylation at two classes of non-imprinted genes. *Dev.* **141**, 1313–1323.

Sagi, I., De Pinho, J. C., Zuccaro, M. V., Atzmon, C., Golan-Lev, T., Yanuka, O., Prosser, R., Sadowy, A., Perez, G., Cabral, T., et al. (2019). Distinct Imprinting Signatures and Biased Differentiation of Human Androgenetic and Parthenogenetic Embryonic Stem Cells. *Cell Stem Cell* **25**, 419-432.e9.

Sanchez-Delgado, M., Riccio, A., Eggerman, T., Maher, E. R., Lapunzina, P., Mackay, D. and Monk, D. (2016a). *Causes and consequences of multi-locus imprinting disturbances in humans.*

Sanchez-Delgado, M., Court, F., Vidal, E., Medrano, J., Monteagudo-Sánchez, A., Martin-Trujillo, A., Tayama, C., Iglesias-Platas, I., Kondova, I., Bontrop, R., et al. (2016b). Human Oocyte-Derived Methylation Differences Persist in the Placenta Revealing Widespread Transient Imprinting. *PLoS Genet.* **12**,.

Sandoval, J., Heyn, H., Moran, S., Serra-Musach, J., Pujana, M. A., Bibikova, M. and Esteller, M. (2011). Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome.

Sanli, I. and Feil, R. (2015). Chromatin mechanisms in the developmental control of imprinted gene expression. *Int. J. Biochem. Cell Biol.* **67**, 139–147.

Sassaman, D. M., Dombroski, B. A., Moran, J. V., Kimberland, M. L., Naas, T. P., DeBerardinis, R. J., Gabriel, A., Swergold, G. D. and Kazazian, H. H. (1997). Many human L1 elements are capable of retrotransposition. *Nat. Genet.* **16**, 37–43.

Schaaf, C. P., Gonzalez-Garay, M. L., Xia, F., Potocki, L., Gripp, K. W., Zhang, B., Peters, B. A., McElwain, M. A., Drmanac, R., Beaudet, A. L., et al. (2013). Truncating mutations of MAGEL2 cause Prader-Willi phenotypes and autism. *Nat. Genet.* **45**, 1405–1409.

Schoenfelder, S. and Fraser, P. (2019). Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455.

Schoenherr, C. J., Levorse, J. M. and Tilghman, S. M. (2003). CTCF maintains differential methylation at the Igf2/H19 locus. *Nat. Genet.* **33**, 66–69.

Schultz, R. M. (1993). Regulation of zygotic gene activation in the mouse. *BioEssays* **15**, 531–538.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T. A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., et al. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908–912.

Sharif, J., Endo, T. A., Nakayama, M., Karimi, M. M., Shimada, M., Katsuyama, K., Goyal, P., Brind'Amour, J., Sun, M.-A., Sun, Z., et al. (2016). Activation of Endogenous Retroviruses in Dnmt1−/− ESCs Involves Disruption of SETDB1-Mediated Repression by NP95 Binding to Hemimethylated DNA. *Cell Stem Cell* **19**, 81–94.

Sharp, A. J., Migliavacca, E., Dupre, Y., Stathaki, E., Sailani, M. R., Baumer, A., Schinzel, A., Mackay, D. J., Robinson, D. O., Cobellis, G., et al. (2010). Methylation profiling in individuals with uniparental disomy identifies novel differentially methylated regions on chromosome 15. *Genome Res.* **20**, 1271–8.

Shemer, R., Birger, Y., Riggs, A. D. and Razin, A. (1997). Structure of the imprinted mouse Snrpn gene and establishment of its parental-specific methylation pattern. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10267–72.

Shenker, N. S., Ueland, P. M., Polidoro, S., van Veldhoven, K., Ricceri, F., Brown, R., Flanagan, J. M. and Vineis, P. (2013). DNA Methylation as a Long-term Biomarker of Exposure to Tobacco Smoke. *Epidemiology* **24**, 712–716.

Shi, D. Q., Ali, I., Tang, J. and Yang, W. C. (2017). New insights into 5hmC DNA modification: Generation, distribution and function. *Front. Genet.* **8**,.

Shi, H., Strogantsev, R., Takahashi, N., Kazachenka, A., Lorincz, M. C., Hemberger, M. and Ferguson-Smith, A. C. (2019). ZFP57 regulation of transposable elements and gene expression within and beyond imprinted domains. *Epigenetics Chromatin* **12**, 49.

Shirane, K., Toh, H., Kobayashi, H., Miura, F., Chiba, H., Ito, T., Kono, T. and Sasaki, H. (2013). Mouse Oocyte Methylomes at Base Resolution Reveal Genome-Wide Accumulation of Non-CpG Methylation and Role of DNA Methyltransferases. *PLoS Genet.* **9**, e1003439.

Silver, H. K., Kiyasu, W., George, J. and Deamer, W. C. (1953). Syndrome of congenital hemihypertrophy, shortness of stature, and elevated urinary gonadotropins. *Pediatrics* **12**, 368–76.

Simpson, A. J. G., Caballero, O. L., Jungbluth, A., Chen, Y. T. and Old, L. J. (2005). Cancer/testis antigens, gametogenesis and cancer. *Nat. Rev. Cancer* **5**, 615–625.

Singh, P., Lee, D. H. and Szabó, P. E. (2012). More than insulator: Multiple roles of CTCF at the H19-Igf2 imprinted domain. *Front. Genet.* **3**,.

Smilinich, N. J., Day, C. D., Fitzpatrick, G. V., Caldwell, G. M., Lossie, A. C., Cooper, P. R., Smallwood, A. C., Joyce, J. A., Schofield, P. N., Reik, W., et al. (1999). A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 8064–8069.

Smit, A. F. A. and Riggs, A. D. (1996). Tiggers and other DNA transposon fossils in the human genome. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1443–1448.

Smith, Z. D. and Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204–220.

Song, J., Rechkoblit, O., Bestor, T. H. and Patel, D. J. (2011). Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science (80-.).* **331**, 1036–1040.

Song, J., Teplova, M., Ishibe-Murakami, S. and Patel, D. J. (2012). Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. *Science (80-.).* **335**, 709–712.

Sparago, A., Verma, A., Patricelli, M. G., Pignata, L., Russo, S., Calzari, L., De Francesco, N., Del Prete, R., Palumbo, O., Carella, M., et al. (2019). The phenotypic variations of multi-locus imprinting disturbances associated with maternal-effect variants of NLRP5 range from overt imprinting disorder to apparently healthy phenotype. *Clin. Epigenetics* **11**, 190.

Spence, J. E., Perciaccante, R. G., Greig, G. M., Willard, H. F., Ledbetter, D. H., Fielding Hejtmancik, J., Pollack, M. S., O'brien, W. E. and Beaudet, A. L. (1988). *Uniparental Disomy as a Mechanism for Human Genetic Disease*.

Spiteri, B. S., Stafrace, Y. and Calleja-Agius, J. (2017). Silver-Russell syndrome: A review. *Neonatal Netw.* **36**, 206–212.

Srivastava, M., Frolova, E., Rottinghaus, B., Boe, S. P., Grinberg, A., Lee, E., Love, P. E. and Pfeifer, K. (2003). Imprint control element-mediated secondary methylation imprints at the Igf2/H19 locus. *J. Biol. Chem.* **278**, 5977–83.

Stadtfeld, M., Apostolou, E., Ferrari, F., Choi, J., Walsh, R. M., Chen, T., Ooi, S. S. K., Kim, S. Y., Bestor, T. H., Shioda, T., et al. (2012). Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of allg-iPS cell mice from terminally differentiated B cells. *Nat. Genet.* **44**, 398–405.

Stöger, R., Kubička, P., Liu, C.-G. G., Kafri, T., Razin, A., Cedar, H. and Barlow, D. P. P. (1993). Maternal-specific methylation of the imprinted mouse Igf2r locus identifies the

expressed locus as carrying the imprinting signal. *Cell* **73**, 61–71.

Stouder, C., Somm, E. and Paoloni-Giacobino, A. (2011). Prenatal exposure to ethanol: A specific effect on the H19 gene in sperm. *Reprod. Toxicol.* **31**, 507–512.

Strain, L., Warner, J. P., Johnston, T. and Bonthron, D. T. (1995). A human parthenogenetic chimaera. *Nat. Genet.* **11**, 164–169.

Strogantsev, R., Krueger, F., Yamazawa, K., Shi, H., Gould, P., Goldman-Roberts, M., McEwen, K., Sun, B., Pedersen, R. and Ferguson-Smith, A. C. (2015). Allele-specific binding of ZFP57 in the epigenetic regulation of imprinted and non-imprinted monoallelic expression. *Genome Biol.* **16**, 112.

Suetake, I., Mishima, Y., Kimura, H., Lee, Y. H., Goto, Y., Takeshima, H., Ikegami, T. and Tajima, S. (2011). Characterization of DNA-binding activity in the N-terminal domain of the DNA methyltransferase Dnmt3a. *Biochem. J.* **437**, 141–148.

Sun, Z., Wu, Y., Ordog, T., Baheti, S., Nie, J., Duan, X., Hojo, K., Kocher, J. P., Dyck, P. J. and Klein, C. J. (2014). Aberrant signature methylome by DNMT1 hot spot mutation in hereditary sensory and autonomic neuropathy 1E. *Epigenetics* **9**, 1184–1193.

Sur, I. and Taipale, J. (2016). The role of enhancers in cancer. *Nat. Rev. Cancer* **16**, 483–493.

Surani, M. A. H. and Barton, S. C. (1983). Development of gynogenetic eggs in the mouse: Implications for parthenogenetic embryos. *Science (80-.).* **222**, 1034–1036.

Surani, M. A. H., Barton, S. C. and Norris, M. L. (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308**, 548–550.

Szulwach, K. E., Li, X., Li, Y., Song, C. X., Wu, H., Dai, Q., Irier, H., Upadhyay, A. K., Gearing, M., Levey, A. I., et al. (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat. Neurosci.* **14**, 1607–1616.

Tafaj, O. and Jüppner, H. (2017). Pseudohypoparathyroidism: one gene, several syndromes. *J. Endocrinol. Invest.* **40**, 347–356.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., et al. (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science (80-.).* **324**, 930–935.

Takagi, N., Sugawara, O. and Sasaki, M. (1982). Regional and temporal changes in the pattern of X-chromosome replication during the early post-implantation development of the female mouse. *Chromosoma* **85**, 275–86.

Takahashi, N., Coluccio, A., Thorball, C. W., Planet, E., Shi, H., Offner, S., Turelli, P., Imbeault, M., Ferguson-Smith, A. C. and Trono, D. (2019). ZNF445 is a primary regulator of genomic imprinting. *Genes Dev.* **33**, 49–54.

Tam, P. P. L. and Behringer, R. R. (1997). Mouse gastrulation: The formation of a mammalian body plan. *Mech. Dev.* **68**, 3–25.

Tatton-Brown, K., Seal, S., Ruark, E., Harmer, J., Ramsay, E., Del Vecchio Duarte, S., Zachariou, A., Hanks, S., O'Brien, E., Aksglaede, L., et al. (2014). Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. *Nat. Genet.* **46**, 385–388.

Temple, I. K., Cockwell, A., Hassold, T., Pettay, D. and Jacobs, P. (1991). *Maternal uniparental disomy for chromosome 14*.

Thakur, A., Mackin, S. J. S.-J., Irwin, R. E. R. E., O'Neill, K. M. K. M., Pollin, G. and Walsh, C. (2016). Widespread recovery of methylation at gametic imprints in hypomethylated mouse stem cells following rescue with DNMT3A2. **9**, 1–15.

Thürmann, L., Grützmann, K., Klös, M., Bieg, M., Winter, M., Polte, T., Bauer, T., Schick Dipl-

Ing, M., Bewerunge-Hudler, M., Röder, S., et al. (2018). Early-onset childhood atopic dermatitis is related to NLRP2 repression. *J. Allergy Clin. Immunol.* **141**, 1482-1485.e16.

Thursby, S.-J., K Lobo, D., Pentieva, K., Zhang, S.-D., Irwin, R. E. and Walsh, C. P. (2020). CandiMeth: Powerful yet simple visualization and quantification of DNA methylation at candidate genes. *Gigascience* **9**,.

Tie, C. H., Fernandes, L., Conde, L., Robbez-Masson, L., Sumner, R. P., Peacock, T., Rodriguez-Plata, M. T., Mickute, G., Gifford, R., Towers, G. J., et al. (2018). KAP 1 regulates endogenous retroviruses in adult human cells and contributes to innate immune control . *EMBO Rep.* **19**,.

Tong, Z. Bin, Gold, L., Pfeifer, K. E., Dorward, H., Lee, E., Bondy, C. A., Dean, J. and Nelson, L. M. (2000). Mater, a maternal effect gene required for early embryonic development in mice. *Nat. Genet.* **26**, 267–268.

Touati, A., Errea-Dorronsoro, J., Nouri, S., Halleb, Y., Pereda, A., Mahdhaoui, N., Ghith, A., Saad, A., Perez de Nanclares, G. and H'mida ben brahim, D. (2019). Transient neonatal diabetes mellitus and hypomethylation at additional imprinted loci: novel ZFP57 mutation and review on the literature. *Acta Diabetol.* **56**, 301–307.

Travers, M. E., Mackay, D. J. G., Nitert, M. D., Morris, A. P., Lindgren, C. M., Berry, A., Johnson, P. R., Hanley, N., Groop, L. C., McCarthy, M. I., et al. (2013). Insights into the molecular mechanism for type 2 diabetes susceptibility at the KCNQ1 locus from temporal changes in imprinting status in human islets. *Diabetes* **62**, 987–992.

Tsai, H. C., Li, H., Van Neste, L., Cai, Y., Robert, C., Rassool, F. V., Shin, J. J., Harbom, K. M., Beaty, R., Pappou, E., et al. (2012). Transient Low Doses of DNA-Demethylating Agents Exert Durable Antitumor Effects on Hematological and Epithelial Tumor Cells. *Cancer Cell* **21**, 430–446.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S. I., Sakaue, M., Matsuoka, C.,

Shimotohno, K., Ishikawa, F., Li, E., Ueda, H. R., et al. (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes to Cells* **11**, 805–814.

Unoki, M. (2019). Recent Insights into the Mechanisms of De Novo and Maintenance of DNA Methylation in Mammals . In *DNA Methylation Mechanism*, p. IntechOpen.

Urbach, A. and Benvenisty, N. (2009). Studying early lethality of 45,XO (Turner's syndrome) embryos using human embryonic stem cells. *PLoS One* **4**,.

Valente, F. M., Sparago, A., Freschi, A., Hill-Harfe, K., Maas, S. M., Frints, S. G. M., Alders, M., Pignata, L., Franzese, M., Angelini, C., et al. (2019). Transcription alterations of KCNQ1 associated with imprinted methylation defects in the Beckwith–Wiedemann locus. *Genet. Med.* **21**, 1808–1820.

van de Lagemaat, L. N., Landry, J.-R., Mager, D. L. and Medstrand, P. (2003). Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.* **19**, 530–536.

van Dongen, J., Nivard, M. G., Willemsen, G., Hottenga, J.-J., Helmer, Q., Dolan, C. V., Ehli, E. A., Davies, G. E., van Iterson, M., Breeze, C. E., et al. (2016). Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nat. Commun.* **7**, 11115.

Van Tongelen, A., Loriot, A. and De Smet, C. (2017). Oncogenic roles of DNA hypomethylation through the activation of cancer-germline genes. *Cancer Lett.* **396**, 130–137.

Velasco, G. and Francastel, C. (2019). Genetics meets DNA methylation in rare diseases. *Clin. Genet.* **95**, 210–220.

Velasco, G., Hube, F., Rollin, J., Neuillet, D., Philippe, C., Bouzinba-Segard, H., Galvani, A., Viegas-Pequignot, E. and Francastel, C. (2010). Dnmt3b recruitment through E2F6

transcriptional repressor mediates germ-line gene silencing in murine somatic tissues. *Proc. Natl. Acad. Sci.* **107**, 9281–9286.

Velasco, G., Walton, E. L., Sterlin, D., Hédouin, S., Nitta, H., Ito, Y., Fouyssac, F., Mégarbané, A., Sasaki, H., Picard, C., et al. (2014). Germline genes hypomethylation and expression define a molecular signature in peripheral blood of ICF patients: Implications for diagnosis and etiology. *Orphanet J. Rare Dis.* **9**,.

Vértesy, Á., Arindrarto, W., Roost, M. S., Reinius, B., Torrens-Juaneda, V., Bialecka, M., Moustakas, I., Ariyurek, Y., Kuijk, E., Mei, H., et al. (2018). Parental haplotype-specific single-cell transcriptomics reveal incomplete epigenetic reprogramming in human female germ cells. *Nat. Commun.* **9**, 1–10.

Von Känel, T. and Huber, A. R. (2013). DNA methylation analysis. *Swiss Med. Wkly.* **143**,.

Wakeling, E. L., Brioude, F., Mackay, D. J. G. and Netchine, I. (2016). Diagnosis and management of Silver–Russell syndrome: first international consensus statement Emma. *Justin H. Davies* **12**,.

Walsh, C. P. and Xu, G. L. (2006). Cytosine methylation and DNA repair. *Curr. Top. Microbiol. Immunol.* **301**, 283–315.

Walsh, C. P., Chaillet, J. R. and Bestor, T. H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation [4]. *Nat. Genet.* **20**, 116–117.

Walter, M., Teissandier, A., Pérez-Palacios, R. and Bourc'His, D. (2016). An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *Elife* **5**,.

Wang, H. and Dey, S. K. (2006). Roadmap to embryo implantation: Clues from mouse models. *Nat. Rev. Genet.* **7**, 185–199.

Wang, T., Chen, K., Zeng, X., Yang, J., Wu, Y., Shi, X., Qin, B., Zeng, L., Esteban, M. A., Pan, G., et al. (2011). The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. *Cell Stem Cell* **9**, 575–587.

Wang, S., Zhang, C., Hasson, D., Desai, A., SenBanerjee, S., Magnani, E., Ukomadu, C., Lujambio, A., Bernstein, E. and Sadler, K. C. (2019). Epigenetic Compensation Promotes Liver Regeneration. *Dev. Cell* **50**, 43-56.e6.

Wang, K. H., Kupa, J., Duffy, K. A. and Kalish, J. M. (2020). Diagnosis and Management of Beckwith-Wiedemann Syndrome. *Front. Pediatr.* **7**, 562.

Weber, M., Davies, J. J., Wittig, D., Oakeley, E. J., Haase, M., Lam, W. L. and Schübeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37**, 853–862.

Weber, A. R., Krawczyk, C., Robertson, A. B., Kuśnierczyk, A., Vågbø, C. B., Schuermann, D., Klungland, A. and Schär, P. (2016). Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nat. Commun.* **7**, 10806.

Weksberg, R., Shuman, C. and Beckwith, J. B. (2010). Beckwith-Wiedemann syndrome. *Eur. J. Hum. Genet.* **18**, 8–14.

Wiedemann, H. (1964). Complexe malformatif familial avec hernie ombilicale et macroglossie—un “syndrome nouveau”? *J. Genet. Hum.* **13**, 223–32.

Wiench, M., John, S., Baek, S., Johnson, T. A., Sung, M. H., Escobar, T., Simmons, C. A., Pearce, K. H., Biddie, S. C., Sabo, P. J., et al. (2011). DNA methylation status predicts cell type-specific enhancer activity. *EMBO J.* **30**, 3028–3039.

Wigler, M., Levy, D. and Perucho, M. (1981). The somatic replication of DNA methylation. *Cell* **24**, 33–40.

Williams, C. A., Angelman, H., Clayton-Smith, J., Driscoll, D. J., Hendrickson, J. E., Knoll, J. H. M., Magenis, R. E., Schinzel, A., Wagstaff, J., Whidden, E. M., et al. (1995). Angelman syndrome: Consensus for diagnostic criteria. *Am. J. Med. Genet.* **56**, 237–238.

Williams, C. A., Beaudet, A. L., Clayton-Smith, J., Knoll, J. H., Kyllerman, M., Laan, L. A., Magenis, R. E., Moncla, A., Schinzel, A. A., Summers, J. A., et al. (2006). Angelman syndrome 2005: Updated consensus for diagnostic criteria. *Am. J. Med. Genet.* **140 A**, 413–418.

Wilson, M., Peters, G., Bennetts, B., McGillivray, G., Wu, Z. H., Poon, C. and Algar, E. (2008). The clinical phenotype of mosaicism for genome-wide paternal uniparental disomy: Two new reports. *Am. J. Med. Genet. Part A* **146**, 137–148.

Winkelmann, J., Lin, L., Schormair, B., Kornum, B. R., Faraco, J., Plazzi, G., Melberg, A., Cornelio, F., Urban, A. E., Pizza, F., et al. (2012). Mutations in DNMT1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Hum. Mol. Genet.* **21**, 2205–2210.

Wiznerowicz, M., Jakobsson, J., Szulc, J., Liao, S., Quazzola, A., Beermann, F., Aebischer, P. and Trono, D. (2007). The Krüppel-associated box repressor domain can trigger de novo promoter methylation during mouse early embryogenesis. *J. Biol. Chem.* **282**, 34535–34541.

Woodcock, D. M., Lawler, C. B., Linsenmeyer, M. E., Doherty, J. P. and Warren, W. D. (1997). Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. *J. Biol. Chem.* **272**, 7810–7816.

Woodfine, K., Huddleston, J. E. and Murrell, A. (2011). Quantitative analysis of DNA methylation at all human imprinted regions reveals preservation of epigenetic stability in adult somatic tissue. *Epigenetics Chromatin* **4**, 1.

Wu, H. and Sun, Y. E. (2006). Epigenetic Regulation of Stem Cell Differentiation. *Pediatr. Res.* **59**, 21R-25R.

Wu, H. and Zhang, Y. (2014). Reversing DNA methylation: Mechanisms, genomics, and biological functions. *Cell* **156**, 45–68.

Wu, X. and Zhang, Y. (2017). TET-mediated active DNA demethylation: Mechanism, function and beyond. *Nat. Rev. Genet.* **18**, 517–534.

Wyatt, G. R. and Cohen, S. S. (1952). A New Pyrimidine Base from Bacteriophage Nucleic Acids. *Nature* **170**, 1072–1073.

Xenopoulos, P., Kang, M. and Hadjantonakis, A.-K. (2012). Cell Lineage Allocation Within the Inner Cell Mass of the Mouse Blastocyst.pp. 185–202.

Xia, W., Xu, J., Yu, G., Yao, G., Xu, K., Ma, X., Zhang, N., Liu, B., Li, T., Lin, Z., et al. (2019). Resetting histone modifications during human parental-to-zygotic transition. *Science (80-.).* **365**, 353–360.

Xu, G. L., Bestor, T. H., BourC'His, D., Hsieh, C. L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J. J. and Viegas-Péquignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**, 187–191.

Yakoreva, M., Kahre, T., Žordania, R., Reinson, K., Teek, R., Tillmann, V., Peet, A., Õiglane-Shlik, E., Pajusalu, S., Murumets, Ü., et al. (2019). A retrospective analysis of the prevalence of imprinting disorders in Estonia from 1998 to 2016. *Eur. J. Hum. Genet.* **27**, 1649–1658.

Yamazawa, K., Nakabayashi, K., Kagami, M., Sato, T., Saitoh, S., Horikawa, R., Hizuka, N. and Ogata, T. (2010). Parthenogenetic chimaerism/mosaicism with a Silver-Russell syndrome-like phenotype. *J. Med. Genet.* **47**, 782–785.

Yamazawa, K., Nakabayashi, K., Matsuoka, K., Masubara, K., Hata, K., Horikawa, R. and Ogata, T. (2011). Androgenetic/biparental mosaicism in a girl with Beckwith-Wiedemann

syndrome-like and upd(14)pat-like phenotypes. *J. Hum. Genet.* **56**, 91–93.

Yang, F. and Wang, P. J. (2016). Multiple LINEs of retrotransposon silencing mechanisms in the mammalian germline. *Semin. Cell Dev. Biol.* **59**, 118–125.

Yoder, J. A., Walsh, C. P. and Bestor, T. H. (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**, 335–40.

Yoo, C. B., Jeong, S., Egger, G., Liang, G., Phiasivongsa, P., Tang, C., Redkar, S. and Jones, P. A. (2007). Delivery of 5-aza-2′-deoxycytidine to cells using oligodeoxynucleotides. *Cancer Res.* **67**, 6400–6408.

Yuan, X.-Q., Chen, P., Du, Y.-X., Zhu, K.-W., Zhang, D.-Y., Yan, H., Liu, H., Liu, Y.-L., Cao, S., Zhou, G., et al. (2019). Influence of DNMT3A R882 mutations on AML prognosis determined by the allele ratio in Chinese patients. *J. Transl. Med.* **17**, 220.

Zahova, S. and Isles, A. (2018). The Role of the Prader-Willi Syndrome Critical Interval for Epigenetic Regulation, Transcription and Phenotype. *Epigenomes* **2**, 18.

Zeng, Y. and Chen, T. (2019). DNA methylation reprogramming during mammalian development. *Genes (Basel).* **10**,.

Zhang, P., Dixon, M., Zucchelli, M., Hambiliki, F., Levkov, L., Hovatta, O. and Kere, J. (2008). Expression Analysis of the NLRP Gene Family Suggests a Role in Human Preimplantation Development. *PLoS One* **3**, e2755.

Zhao, Q., Zhang, J., Chen, R., Wang, L., Li, B., Cheng, H., Duan, X., Zhu, H., Wei, W., Li, J., et al. (2016). Dissecting the precise role of H3K9 methylation in crosstalk with DNA maintenance methylation in mammals. *Nat. Commun.* **7**, 1–12.

Zheng, Y., Joyce, B. T., Colicino, E., Liu, L., Zhang, W., Dai, Q., Shrubsole, M. J., Kibbe, W. A., Gao, T., Zhang, Z., et al. (2016a). Blood Epigenetic Age may Predict Cancer Incidence and

Mortality. *EBioMedicine* **5**, 68–73.

Zheng, H., Huang, B., Zhang, B., Xiang, Y., Du, Z., Xu, Q., Li, Y., Wang, Q., Ma, J., Peng, X., et al. (2016b). Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals. *Mol. Cell* **63**, 1066–1079.

Zhou, W., Laird, P. W. and Shen, H. (2017). Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res.* **45**, e22.

Zilberman, D. (2017). An evolutionary case for functional gene body methylation in plants and animals. *Genome Biol.* **18**,.

Ziller, M. J., Müller, F., Liao, J., Zhang, Y., Gu, H., Bock, C., Boyle, P., Epstein, C. B., Bernstein, B. E., Lengauer, T., et al. (2011). Genomic distribution and Inter-Sample variation of Non-CpG methylation across human cell types. *PLoS Genet.* **7**,.